

A POSTERIORI ANALYSIS FOR SPACE-TIME, DISCONTINUOUS IN TIME GALERKIN APPROXIMATIONS FOR PARABOLIC EQUATIONS IN A VARIABLE DOMAIN

DIMITRA ANTONOPOULOU^{2,3,*} AND MICHAEL PLEXOUSA^{KIS^{1,3}}

Abstract. This paper presents an *a posteriori* error analysis for the discontinuous in time space–time scheme proposed by Jamet for the heat equation in multi-dimensional, non-cylindrical domains Jamet (*SIAM J. Numer. Anal.* **15** (1978) 913–928). Using a Clément-type interpolant, we prove abstract *a posteriori* error bounds for the numerical error. Furthermore, in the case of two-dimensional spatial domains we transform the problem into an equivalent one, of parabolic type, with space–time dependent coefficients but posed on a cylindrical domain. We formulate a discontinuous in time space–time scheme and prove *a posteriori* error bounds of optimal order. The *a priori* estimates of Evans (American Mathematical Society (1998)) for general parabolic initial and boundary value problems are used in the derivation of the upper bound. Our lower bound coincides with that of Picasso (*Comput. Meth. Appl. Mech. Eng.* **167** (1998) 223–237), proposed for adaptive, Runge–Kutta finite element methods for linear parabolic problems. Our theoretical results are verified by numerical experiments.

Mathematics Subject Classification. 65M12, 65M15, 65M60.

Received December 28, 2016. Accepted September 25, 2018.

1. INTRODUCTION

1.1. The problem

The discontinuous Galerkin method was introduced by Reed and Hill [37] and Lesaint and Raviart [29]. This method has been proven to be efficient when applied to problems posed on domains of complicated geometry, approximates well discontinuous solutions and can be combined effectively with refinement or adaptivity techniques; see for instance [12–14, 17]. Jamet [25] considered the discontinuous Galerkin method for parabolic problems in general space–time finite element spaces.

The problem considered in this paper is posed on a domain of variable geometry in time and thus involves a moving boundary. We refer to the classical results of Baines and Miller [9, 34] on the development of the so-called moving finite elements for the numerical approximation of a wide class of PDEs of hyperbolic and parabolic type.

Keywords and phrases. Heat equation, space–time discontinuous Galerkin methods, *a posteriori* error estimates, non-cylindrical domains.

¹ Department of Mathematics and Applied Mathematics, University of Crete, 714 09 Heraklion, Greece.

² Department of Mathematics, University of Chester, Thornton Science Park, CH2 4NU, UK.

³ Institute of Applied and Computational Mathematics, FORTH, 711 10 Heraklion, Greece.

*Corresponding author: d.antonopoulou@chester.ac.uk

Makridakis and Nochetto [31] presented an *a posteriori* error analysis for high order discontinuous Galerkin numerical schemes for evolutionary problems. In [2], issues such as optimality of estimates, adaptivity and nodal convergence rates were investigated, while in [26] a discontinuous Galerkin finite element method was applied to the non-linear Schrödinger equation in cylindrical domains. Picasso [36] analyzed Runge-Kutta, adaptive finite element methods for linear parabolic problems.

More recently, Antonopoulou and Plexousakis [7], presented an *a priori* error analysis for a method analogous to (1.3) for the linear Schrödinger equation in non-cylindrical domains; see also the interesting work of Schötzau and Wihler [38] for the *a posteriori* error analysis of hp-methods for the time discretization of parabolic problems with continuous and discontinuous Galerkin methods.

The numerical scheme proposed in this work is a genuine space-time finite element method. Recently, there has been a renewed interest on high order space-time schemes due to their effectiveness when applied, for example, to problems in fluid dynamics, elasticity or conservation laws, even on unstructured grids. We refer to the works of van der Vegt and van der Ven on inviscid compressible flows [41, 42]; see also [28] for the case of the compressible Navier–Stokes equations, and the recent results of Dumbser and Tavelli for three-dimensional Navier–Stokes equations and linear elasticity for finite elements on unstructured meshes [39, 40]. May and Zakerzadeh [32] proved convergence of space-time discontinuous Galerkin approximations for scalar hyperbolic conservation laws by avoiding the so-called streamline–diffusion stabilization. Hiltebrand, Mishra and Parés constructed entropy–stable approximations for non-conservative hyperbolic systems [22], while in [21], the authors considered multi-dimensional nonlinear systems of conservation laws.

We consider a bounded domain $\Omega(t)$ in R^m , $m \geq 1$, continuously dependent on $t \in [0, T]$. We let $\Gamma(t)$ denote the boundary of $\Omega(t)$ and

$$S_T := \{(x, t) : x \in \Omega(t), 0 < t < T\}, \quad \Sigma_T := \{(x, t) : x \in \Gamma(t), 0 < t < T\}.$$

We seek $u : \overline{S}_T \rightarrow R$ such that

$$\begin{aligned} u_t &= \Delta u + f & \text{in } S_T, \\ u &= 0 & \text{in } \Sigma_T, \\ u &= u^0 & \text{in } \Omega(0), \end{aligned} \tag{1.1}$$

where f is a function defined on S_T and u^0 on $\Omega(0)$. Detailed *a priori* error analysis of discontinuous in time space-time schemes for this problem has been carried out by Jamet [25]. We note that (1.1) is posed on a domain varying in time, so it may be thought of as a Stefan type problem. However, the boundary $\Gamma(t)$ does not depend on the solution. Hence, it is not a free boundary problem.

We employ standard notation for the Sobolev spaces used throughout. For S a subdomain of S_T , we denote by $\tilde{H}^1(S)$ the space of functions in $H^1(S)$ that vanish on $\Sigma_T \cap \overline{S}$. We let $((\cdot, \cdot))_S$ denote the inner product and $\|\cdot\|_S$ the corresponding norm in $L^2(S)$. Similarly, we denote by $(\cdot, \cdot)_{\Omega(t)}$ the inner product in $L^2(\Omega(t))$ and by $|\cdot|_{\Omega(t)}$ the corresponding norm.

For $0 = t^0 < t^1 < \dots < t^N = T$, a partition of $[0, T]$, we write $\Omega^n := \Omega(t^n)$ and

$$G^n := \{(x, t) : x \in \Omega(t), t \in (t^n, t^{n+1})\}, \quad \tilde{G}^n := \{(x, t) : x \in \overline{\Omega}(t), t \in (t^n, t^{n+1})\}.$$

Also, for $0 \leq \tau_0 < \tau_1 \leq T$, we let

$$G(\tau_0, \tau_1) = \{(x, t) : x \in \Omega(t), \tau_0 < t < \tau_1\}.$$

For each $0 \leq n \leq N - 1$ we consider a family $\{V_h^n\}$ of finite dimensional subspaces of $\tilde{H}^1(G^n)$ parametrized by $0 < h \leq 1$. V_h will denote the space of all functions defined on \overline{S}_T such that their restriction on each \tilde{G}^n coincides with the restriction to \tilde{G}^n of a function $v_h \in V_h^n$. Functions in V_h are, in general, discontinuous at the temporal nodes t^n . Consequently, we shall use the notation introduced in [25], and also used in [4, 7],

$$v_h^n(\cdot) := v_h(\cdot, t^n) \quad \text{for } 0 \leq n \leq N,$$

and

$$v_h^{n+0}(\cdot) := \lim_{\epsilon \rightarrow 0^+} v_h(\cdot, t^n + \epsilon) \quad \text{for } 0 \leq n \leq N-1.$$

Here, note that $v_h(\cdot, t^n) = \lim_{\epsilon \rightarrow 0^+} v_h(\cdot, t^n - \epsilon)$.

In the analysis below we shall employ a transformation of the form $u \rightarrow e^{ct}u$, with c a suitable constant, and consider the initial and boundary value problem

$$\begin{aligned} u_t &= \Delta u + \beta u + f & \text{in } S_T, \\ u &= 0 & \text{in } \Sigma_T, \\ u &= u^0 & \text{in } \Omega(0), \end{aligned} \tag{1.2}$$

where β is a suitably small, negative constant.

Remark 1.1. The necessity of this transformation will be analyzed in the sequel. We shall show that it enables us to, for example, use an L^2 -stability argument in the proof of the upper bound (2.7), and derive the *a posteriori* L^2 norm estimate (2.23) for the initial problem posed on the non-cylindrical domain. Moreover, when our scheme is applied to the equivalent continuous problem, a second order parabolic equation posed on a cylindrical domain, with space-time dependent coefficients, we prove existence of the numerical solution by L^2 -stability, *cf.* Theorem 3.1, and derive an H^1 estimate for the *a posteriori* error. The exponential transformation $u \rightarrow e^{ct}u$ was introduced by Antonopoulou [4] and has been used in [5–7] for the linear Schrödinger equation.

1.2. The numerical scheme

The discontinuous Galerkin method for (1.2) is defined as follows: Find $u_h \in V_h$ satisfying

$$\begin{aligned} B_n(u_h, v_h) &= ((f, v_h))_{G^n}, \quad \forall v_h \in V_h^n, \quad n = 0, \dots, N-1, \\ u_h^0 &= u^0, \end{aligned} \tag{1.3}$$

where the bilinear form $B_n(u_h, v_h)$ is given by

$$\begin{aligned} B_n(u_h, v_h) &:= -((u_h, \partial_t v_h))_{G^n} + ((\nabla u_h, \nabla v_h))_{G^n} \\ &\quad + (u_h^{n+1}, v_h^{n+1})_{\Omega^{n+1}} - (u_h^n, v_h^{n+0})_{\Omega^n} - ((\beta u_h, v_h))_{G^n}. \end{aligned} \tag{1.4}$$

Existence and *a priori* estimates for finite element spaces of general type for the proposed numerical scheme (1.4) have been proved by Jamet [25] for $\beta = 0$. The same results easily extend for nonzero β .

1.3. Main results and estimation strategy

In this work we undertake the *a posteriori* error analysis of the classical space-time numerical scheme of Jamet ([25]) proposed for parabolic problems posed on non-cylindrical domains. In Section 2, we present an *a posteriori* error analysis for problem (1.2). We derive upper bounds for the error by utilizing the Clément type interpolant of Bernardi [11] in the multi-dimensional case. In the analysis we assume a sufficiently smooth space-time boundary and use local trace estimates. Our finite element space consists of piecewise linear functions in the t and x variables. Since $\dim(\Omega(t)) = m$, the inner elements of our partitions consist of $m+1$ -simplices, while the boundary elements are assumed compatible with a possibly curved boundary, [11].

In Section 3, we analyze the two-dimensional case of the spatial variables and transform our problem into an equivalent one posed on a parallelepiped of \mathbb{R}^3 (1-dimension for time and 2-dimensions for space). The new problem is of the general parabolic form, with space-time dependent coefficients. Since the transformed domain is cylindrical, space-time integration is commutative. This enables the coupling of the local L^2 -projection in time with the standard 2-dimensional Clément interpolant in space, [16]. The interpolation error is estimated in various norms. This is achieved by using certain local trace and Sobolev inequalities together with the well known space-time *a priori* high-order estimates of the continuous problem presented in [19]. As a result, we

derive an *a posteriori* upper bound of optimal order. In addition, following a method proposed by Picasso [36] we prove a lower bound of optimal order.

In Section 4, considering the case $m = 2$, *i.e.* when the spatial domain is in \mathbb{R}^2 , we implement the proposed scheme in a FORTRAN/C++ code. The experimental results verify the optimal order of convergence. Furthermore, the *a posteriori* upper bound constants are approximated numerically. Their computed values are small, a fact that motivates strongly the use of the proposed estimator in applications.

Finally, in the Appendix, we present some trace and Sobolev inequalities that we use throughout our analysis. Here, we focus on the dependence of the estimates on the diameter of the domain, so as to apply them locally, for each element of the partition. In this case the bounds will involve the discretization parameters. We note, also, that the finite element partitions considered in this work are assumed to be regular, uniformly on n , but we do not impose any other restriction on the meshes. Two main ideas, related to the derivation of the *a posteriori* upper bounds, are presented and implemented in this work:

- (1) *The use of the regularity of the continuous problem.* The error $\varepsilon := u - u_h$ of the numerical scheme is the difference between the continuous and the discrete solution. In order to derive upper bounds, since our method uses a space-time finite element formulation, we construct an appropriate interpolant of ε in space and time variables. As a result, high order derivatives of the error ε appear in the upper bound, stemming from the interpolation error, which cannot be estimated directly. Restricting ourselves to each element of the partition, we estimate the derivatives of the exact solution by using the space-time *a priori* high order estimates of Evans [19] for the general parabolic initial and boundary value problems with space-time coefficients. Since the bound is *a posteriori*, the derivatives of the numerical solution may be used in the estimator.
- (2) *We control the constants appearing in the estimator and the error by choosing sufficiently high order polynomial approximations in time. This choice affects only slightly the estimator's computational cost compared to the alternate choice of a high order approximation in space, which is multi-dimensional.* This flexibility is seen as an advantage of our method since it allows its implementation without having to change the numerical scheme at each time interval; the degree of polynomial approximation in time can be independently chosen, as high as we wish, between two temporal nodes since the scheme is discontinuous in time, in contrast to, say, Runge-Kutta (RK) methods. Considering high order RK methods, we refer to the classical result of Alexander, who introduced the diagonally implicit RK methods (DIRK) for stiff ODEs, [3]; *cf.* also in [18], or in [35] for the application of DIRK schemes to hyperbolic systems.

2. A POSTERIORI ERROR ANALYSIS ON THE INITIAL DOMAIN

2.1. Basic identities

Let us define, for $x \in \Omega(t)$,

$$u^n(x) := \lim_{\epsilon \rightarrow 0^+} u(x, t^n - \epsilon) \quad \text{for } 0 \leq n \leq N,$$

and

$$u^{n+0}(x) := \lim_{\epsilon \rightarrow 0^+} u(x, t^n + \epsilon) \quad \text{for } 0 \leq n \leq N - 1,$$

where u is the solution of (1.2). Obviously, if u is continuous on t^n , then $u^n(x) = u^{n+0}(x) = u(x, t^n)$. If $u \in L^2(G^n)$, $\|\nabla u\|_{G^n} < \infty$, and $u \in L^2(\Omega^n) \cap L^2(\Omega^{n+1})$, for any $0 \leq n \leq N - 1$, then for any $v_h \in V_h$ the following equality holds true

$$\begin{aligned} ((f, v_h))_{G^n} &= -((u, \partial_t v_h))_{G^n} + ((\nabla u, \nabla v_h))_{G^n} \\ &\quad + (u, v_h^{n+1})_{\Omega^{n+1}} - (u, v_h^{n+0})_{\Omega^n} - ((\beta u, v_h))_{G^n} \\ &= -((u, \partial_t v_h))_{G^n} + ((\nabla u, \nabla v_h))_{G^n} \\ &\quad + (u^{n+1}, v_h^{n+1})_{\Omega^{n+1}} - (u^{n+0}, v_h^{n+0})_{\Omega^n} - ((\beta u, v_h))_{G^n}, \end{aligned} \tag{2.1}$$

for any $0 \leq n \leq N - 1$. We define the error $\varepsilon(x, t) := u(x, t) - u_h(x)$, and set

$$\varepsilon^n(x) := \lim_{\epsilon \rightarrow 0^+} u(x, t^n - \epsilon) - u_h^n(x), \quad x \in \Omega(t).$$

Remark 2.1. We recall, [25] that a weak solution of (1.2) is defined as satisfying the weak formulation

$$\begin{aligned} ((f, \phi))_{G^n} = & - ((u, \partial_t \phi))_{G^n} + ((\nabla u, \nabla \phi))_{G^n} \\ & + (u, \phi)_{\Omega^{n+1}} - (u, \phi)_{\Omega^n} - ((\beta u, \phi))_{G^n}, \end{aligned} \quad (2.2)$$

for any Lipschitz continuous ϕ defined on $\overline{G^n}$ and $0 \leq n \leq N - 1$. Clearly, if the solution u of (1.2) satisfies, for any $0 \leq n \leq N - 1$,

$$u \in L^2(G^n), \quad \|\nabla u\|_{G^n} < \infty, \quad u \in L^2(\Omega^n) \cap L^2(\Omega^{n+1}), \quad f \in L^2(G(0, T)),$$

then it is a weak solution in the above sense. See [4, 7] for an analogous argument for the linear Schrödinger equation. Of course, u may be more regular. If $f \in L^2(G(0, T))$, $u^0 \in L^2(\Omega(0))$, then a weak solution to (1.2) exists, [25, 30, 33]. Under additional regularity assumptions on f and u^0 , a strong solution u exists, cf. [20, 23, 25], and in this case u is continuous in t , and $u \in H^1(G(0, T))$, where

$$\|u\|_{H^1(G(0, T))} = \left(\|u\|_{L^2(G(0, T))}^2 + \|\nabla u\|_{L^2(G(0, T))}^2 + \|u_t\|_{L^2(G(0, T))}^2 \right)^{1/2}.$$

Our *a posteriori* error analysis in this work, for general $\Omega(t)$, assumes that u is at least continuous in time for any $t \in [0, t^N]$, and that $u \in H^1(G(0, t^N))$, where $t^N = T$, as in [25] where the numerical scheme was proposed. Moreover, the solution $u_h(x, t)$ of (1.3), satisfies $u_h^n(x) \simeq \lim_{\epsilon \rightarrow 0^+} u(x, t^n - \epsilon)$ and hence, $u_h^N(x) \simeq \lim_{\epsilon \rightarrow 0^+} u(x, t^N - \epsilon)$.

Remark 2.2. As we shall see later, in a case of interest where $\Omega(t)$ is in \mathbb{R}^2 , the lateral boundary $\Gamma(t)$ of $\Omega(t)$ consists of a horizontal plane and a surface $s = s(t, \theta)$, a suitable change of variables leads to a second order parabolic equation with an s -dependent initial condition and space-time coefficients involving s, s_t , and up to second order derivatives of s in space (cf. (2.16)). The regularity of the solution of the equivalent parabolic equation depends on the regularity of the initial condition and the regularity of the coefficients. In this case we assume (2.21) for u^0 , f and s , up to $n = N$, which yields the continuity of u and subsequently that $u \in H^1(G(0, T))$.

Remark 2.3. Any loss of continuity or of smoothness of $\Omega(t)$ and of the lateral boundary $\Gamma(t)$ for some t , may give rise to a non smooth or discontinuous solution u at t . Consider, for example, problem (2.16), where such “critical” points t appear when s is discontinuous (or non smooth) there, specifically when s does not satisfy (2.21). These points are, nevertheless, computable since the surface s is known. Our numerical scheme is, of course, applicable up to the first critical point and the relevant error analysis is valid.

In contrast, (1.2) or (2.16), is a linear, parabolic initial and boundary value problem, and thus, any loss of regularity is induced by the initial data. So, if t' is a “critical” point where, for example, (2.21) is not valid, we consider the time interval (t, t') and define as initial time the time t . We then apply again our scheme, but now with initial condition *not* the exact solution $u^N := \lim_{\epsilon \rightarrow 0^+} u(\cdot, t - \epsilon)$, as we did in $(0, t)$ but the already computed approximation $u_h^N \simeq \lim_{\epsilon \rightarrow 0^+} u(\cdot, t - \epsilon)$.

The exact solution u is smooth on any interval of the form $(\tau - \epsilon, \tau + \epsilon) \subset (0, t)$, for any $\tau < t$, therefore, u^N exists and is regular, while any loss of regularity of $u(\tau)$ for $\tau > t$ is due to the low regularity of the coefficients of the pde, involving $s(\tau, \cdot)$ and its derivatives, on intervals of the form $(t - \epsilon, t + \epsilon)$. In this case, a continuous dependence of the solution $u(\cdot, \tau)$ on the exact initial condition u^N for $\tau > t$ near t , is not expected. Moreover, the linear system providing the numerical solution at an interval of the form $(t_{\text{new}}^0 := t = t^N, t_{\text{new}}^1)$ may change

significantly, mainly due to possible jumps of s and its derivatives of order $\mathcal{O}(1)$ at t^N , while the contribution of the initial condition jump $u_h^N - u^N$ in the system is controlled by the discretization parameters. Specifically, up to time t' the schemes proposed in this work are again well posed and the *a priori* and *a posteriori* error estimates, involve an additional term from the non zero difference

$$\varepsilon_{\text{new}}^0 := \varepsilon^N = \varepsilon(\cdot, t^N) = \lim_{\epsilon \rightarrow 0^+} u(\cdot, t^N - \epsilon) - u_h^N,$$

see, *e.g.* (2.4), which is the term $\frac{1}{2}|\varepsilon^N|_{\Omega^N}^2 \neq 0$; this term however is estimated by the *a posteriori* error on the interval $(0, t)$ and it is of the same order as the error in $(0, t)$, see (2.23), or Theorem 3.6. Thus, our analysis is applicable under low regularity assumptions for u , when the problem (1.2) is posed on S_T , even discontinuity, on a *finite subset* of $(0, T)$, and the *a posteriori* estimates hold true. The case of blow-up for s and its derivatives in finite time in $(0, T)$ is not covered, since we need at least L^∞ smoothness in space and time variables, piecewise in $(0, T)$, (see *e.g.* (2.21)); see also the coefficients for the more general problem (3.1).

In what follows and for the rest of this paper, the initial condition for our numerical scheme is the exact initial condition of the continuous problem, *i.e.* $u_h^0 = u^0$. However, in order to address the more general case where $u_h^0 \neq u^0$, which, as stated in the previous remark, may give rise to a non smooth solution u , we retain the initial error term ε^0 in our estimates. Finally, we shall use c and C to denote generic constants, independent of the discretization parameters.

Using (1.3), (1.4), (2.1) and the continuity of u in time, we see that the error $\varepsilon := u - u_h$ satisfies, for any $0 \leq i \leq n-1$,

$$\begin{aligned} & -((\beta\varepsilon, \varepsilon))_{G^i} - ((\varepsilon, \partial_t \varepsilon))_{G^i} + ((\nabla \varepsilon, \nabla \varepsilon))_{G^i} + (\varepsilon^{i+1}, \varepsilon^{i+1})_{\Omega^{i+1}} - (\varepsilon^i, \varepsilon^{i+0})_{\Omega^i} \\ & = -((\varepsilon, \partial_t(u - v_h)))_{G^i} + ((\nabla \varepsilon, \nabla(u - v_h)))_{G^i} \\ & \quad + (\varepsilon^{i+1}, u^{i+1} - v_h^{i+1})_{\Omega^{i+1}} - (\varepsilon^i, u^i - v_h^{i+0})_{\Omega^i} - ((\beta\varepsilon, u - v_h))_{G^i}. \end{aligned} \quad (2.3)$$

Let us define, as in [27], $\eta := \varepsilon - v_h$ and $\eta^n := \varepsilon^n - v_h^n$. We note that

$$\eta^i + u_h^i = u^i - v_h^i, \quad v_h^i = u^i - \eta^i - u_h^i, \quad v_h^{i+0} = u^{i+0} - \eta^{i+0} - u_h^{i+0}.$$

Using again the continuity of u in time and (2.3) we obtain, for any $n \leq N$,

$$\begin{aligned} & -\sum_{i=0}^{n-1} ((\beta\varepsilon, \varepsilon))_{G^i} + \|\nabla \varepsilon\|_{G(0, t^n)}^2 + \frac{1}{2}|\varepsilon^n|_{\Omega^n}^2 + \frac{1}{2}\sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\Omega^i}^2 \\ & = \frac{1}{2}|\varepsilon^0|_{\Omega^0}^2 + ((f, \eta))_{G(0, t^n)} + \sum_{i=0}^{n-1} ((u_h, \partial_t \eta))_{G^i} - \sum_{i=0}^{n-1} ((\nabla u_h, \nabla \eta))_{G^i} \\ & \quad + (u_h^0, \eta^0)_{\Omega^0} - (u_h^n, \eta^n)_{\Omega^n} - \sum_{i=0}^{n-1} (u_h^i, \eta^i - \eta^{i+0})_{\Omega^i} + \sum_{i=0}^{n-1} ((\beta u_h, \eta))_{G^i}, \end{aligned} \quad (2.4)$$

where $\varepsilon^0 = 0$ if $u_h^0 = u^0$. Let \mathcal{T}_h^n be a partition of $\overline{G^n}$ and

$$V_h^n = \{z_h \in \widetilde{H}^1(G^n) : z_h|_K \in P_{\rho-1}(K), \forall K \in \mathcal{T}_h^n\},$$

where $P_{\rho-1}$ is the space of polynomials of total degree at most $\rho-1$ in the time and space variables. We let h_n denote the maximum element diameter in the partition \mathcal{T}_h^n and define $h := \max_n h_n$. Further, if ℓ is an interior edge of \mathcal{T}_h^n we let

$$[\nabla u_h \cdot \mathbf{n}]_\ell := \nabla u_h \cdot \mathbf{n}|_{\ell^+} - \nabla u_h \cdot \mathbf{n}|_{\ell^-}$$

denote the jump of $\nabla u_h \cdot \mathbf{n}$ across the edge ℓ , where \mathbf{n} is the normal direction. We also denote by E_K^n the set of the edges of an element $K \in \mathcal{T}_h^n$.

Theorem 2.4. *If u is the solution of (1.2) and u_h the solution of (1.3), then the error $\varepsilon = u - u_h$ satisfies*

$$\begin{aligned} & - \sum_{i=0}^{n-1} ((\beta \varepsilon, \varepsilon))_{G^i} + \|\nabla \varepsilon\|_{G(0, t^n)}^2 + \frac{1}{2} |\varepsilon^n|_{\Omega^n}^2 + \frac{1}{2} \sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\Omega^i}^2 \\ & = \frac{1}{2} |\varepsilon^0|_{\Omega^0}^2 + \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} ((f - \partial_t u_h + \Delta u_h + \beta u_h, \eta))_K \\ & + \sum_{i=0}^{n-1} (u_h^i - u_h^{i+0}, \eta^{i+0})_{\Omega^i} - \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \int_{\ell} \eta [\nabla u_h \cdot \mathbf{n}]_{\ell} ds, \end{aligned} \quad (2.5)$$

for any $v_h \in V_h$, and for any $n \leq N$, where $\eta = \varepsilon - v_h$ and $E_{K, \text{in}}^i$ is the set of interior edges of an element K of the partition \mathcal{T}_h^i , and $\varepsilon^0 = 0$ if $u_h^0 := u^0$.

Proof. Using integration by parts we have

$$\begin{aligned} ((u_h, \partial_t \eta))_{G^i} & = -((\partial_t u_h, \eta))_{G^i} + (u_h^{i+1}, \eta^{i+1})_{\Omega^{i+1}} - (u_h^{i+0}, \eta^{i+0})_{\Omega^i} \\ & = -((\partial_t u_h, \eta))_{G^i} + (u_h^{i+1}, \eta^{i+1})_{\Omega^{i+1}} - (u_h^i, \eta^{i+0} - \eta^i)_{\Omega^i} \\ & \quad - (u_h^i, \eta^i)_{\Omega^i} + (u_h^i - u_h^{i+0}, \eta^{i+0})_{\Omega^i}, \end{aligned}$$

and thus

$$\begin{aligned} \sum_{i=0}^{n-1} ((u_h, \partial_t \eta))_{G^i} & = - \sum_{i=0}^{n-1} ((\partial_t u_h, \eta))_{G^i} + (u_h^n, \eta^n)_{\Omega^n} - (u_h^0, \eta^0)_{\Omega^0} \\ & \quad + \sum_{i=0}^{n-1} (u_h^i, \eta^i - \eta^{i+0})_{\Omega^i} + \sum_{i=0}^{n-1} (u_h^i - u_h^{i+0}, \eta^{i+0})_{\Omega^i}. \end{aligned}$$

In addition,

$$\sum_{i=0}^{n-1} ((\nabla u_h, \nabla \eta))_{G^i} = - \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} ((\Delta u_h, \eta))_K + \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \int_{\ell} \eta [\nabla u_h \cdot \mathbf{n}]_{\ell} ds.$$

Here, we used the fact that η satisfies the Dirichlet boundary condition along the lateral boundary of G^i and that η is continuous in space variables. Then, using (2.4), we obtain (2.5). \square

2.2. Upper bound

We take $\rho = 2$ so that our finite element space consists of piecewise linear functions in the t and x variables. Let l_n be the minimum diameter of elements in \mathcal{T}_h^n and $l = \min_n l_n$. Since $\dim(\Omega(t)) = m$ we have $\dim \overline{G^n} = m+1$ and the interior elements of the partition consist of $(m+1)$ -simplices, while the boundary elements are assumed to be compatible with a possibly curved boundary, in the sense of Bernardi [11]. We shall assume that there exists a positive constant c_0 such that $h/l \leq c_0$, meaning that the partition is regular, uniformly in n . To derive an upper bound for the *a posteriori* error we shall choose v_h in Theorem 2.4 appropriately: For $n \geq 1$, we define $v_h|_{G^n}$ restricted in every element K of \mathcal{T}_h^n as the Clément's interpolant $\pi_h^n \varepsilon$ of the error ε in $P_1(K)$, [11, 16]. We recall [10, 11, 16] that there exist positive constants, depending only on c_0 , such that,

$$\begin{aligned} \|\varepsilon - \pi_h^n \varepsilon\|_K & \leq Ch \|\varepsilon\|_{1, \Delta K}, \\ \|\varepsilon - \pi_h^n \varepsilon\|_{\ell} & \leq Ch^{1/2} \|\varepsilon\|_{1, \Delta K}, \end{aligned} \quad (2.6)$$

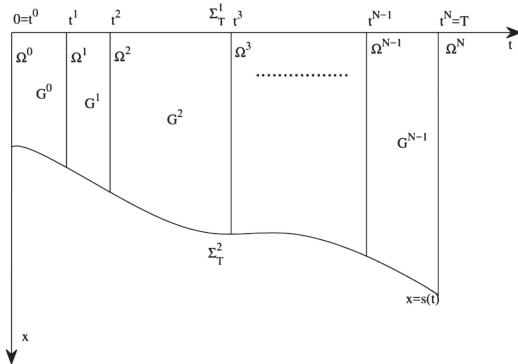


FIGURE 1. The space-time domain in the case $m = \dim(\Omega(t)) = 1$, $\Omega(t) = (0, s(t))$, $\Gamma(t) = \{0, s(t)\}$, $t \in (0, T)$ and $\Sigma_T = \Sigma_T^1 \cup \Sigma_T^2$.

where ℓ is an edge of K and ΔK denotes the set of elements having an edge or vertex common with K . Here, $\|\cdot\|_{\mathcal{A}}$, $\|\cdot\|_{1, \Delta K}$ are the $L^2(\mathcal{A})$ and $H^1(\Delta K)$ norms, respectively, in t and x variables for $\mathcal{A} = K, \ell$. In particular, we note that

$$\|\varepsilon\|_{1, \Delta K} := \left(\|\varepsilon\|_{\Delta K}^2 + \|\partial_t \varepsilon\|_{\Delta K}^2 + \|\nabla \varepsilon\|_{\Delta K}^2 \right)^{1/2},$$

where $\nabla := \nabla_x$. In the next proposition we use the estimates for the interpolant of Clément to establish an upper bound for the error.

Proposition 2.5. *Let $m \geq 2$. If u is the solution of (1.2), u_h the solution of (1.3) and $\varepsilon = u - u_h$, then for any $0 < \epsilon \leq 1$, there exist positive constants c, c_1, c_2, c_3 such that*

$$\begin{aligned} & \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\nabla \varepsilon\|_K^2 + c_1 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_K^2 + c_2 |\varepsilon^n|_{\Omega^n}^2 + c_3 \sum_{i=0}^{n-1} \|u_h^{i+0} - u_h^i\|_{\Omega^i}^2 \\ & \leq c |\varepsilon^0|_{\Omega^0}^2 + ch^{2/(1+\epsilon)} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|f - \partial_t u_h + \Delta u_h + \beta u_h\|_{L^\vartheta(K)} \\ & \quad + ch^2 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|f - \partial_t u_h + \Delta u_h + \beta u_h\|_{L^\vartheta(K)}^\vartheta \\ & \quad + ch^{1/(1+\epsilon)} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \left(\int_\ell |[\nabla u_h \cdot \mathbf{n}]_\ell|^\vartheta ds \right)^{1/\vartheta} \\ & \quad + ch \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \int_\ell |[\nabla u_h \cdot \mathbf{n}]_\ell|^\vartheta ds \\ & \quad + c_0 h^{1-\epsilon} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\partial_t \varepsilon\|_K^2, \end{aligned} \tag{2.7}$$

provided that h is sufficiently small. Here, $\vartheta := (1+\epsilon)/\epsilon$ and c_0 a positive constant, as small as needed. Moreover, $\varepsilon^0 = 0$, if $u_h^0 := u^0$.

Proof. In view of equation (2.5), we shall need estimates for the quantities $|\eta^{i+0}|_{\Omega^i}$, and $\|\eta\|_\ell$ for $\ell \in E_{K,\text{in}}^i$. To this end, using the second estimate of (2.6) for $\ell \in E_{K,\text{in}}^i$ we obtain

$$\left(\int_\ell |\eta|^2 ds \right)^{1/2} = \left(\int_\ell |\varepsilon - \pi_h^i \varepsilon|^2 ds \right)^{1/2} \leq c \left(h \|\varepsilon\|_{1,\Delta K}^2 \right)^{1/2} = c \sqrt{h} \|\varepsilon\|_{1,\Delta K}. \quad (2.8)$$

Now, observe that $\Omega^i \subset \partial G^i, \partial G^{i-1}$, so that

$$|\eta^{i+0}|_{\Omega^i}^2 = \int_{\Omega^i} |\eta(x, t^{i+0})|^2 dx \leq \int_{\partial G^i} |\eta(x, t)|^2 ds.$$

By the local trace inequality (A.1) we arrive at

$$|\eta^{i+0}|_{\Omega^i}^2 \leq c \int_{\partial G^i} |\eta(x, t)|^2 ds \leq c \|\eta\|_{G^i} \|\eta\|_{1,G^i} + ch^{-1} \|\eta\|_{G^i}^2. \quad (2.9)$$

Note that the same argument gives, for $i \geq 1$,

$$|\eta^i|_{\Omega^i}^2 \leq \int_{\partial G^i} |\eta(x, t)|^2 ds \leq c \|\eta\|_{G^{i-1}} \|\eta\|_{1,G^{i-1}} + ch^{-1} \|\eta\|_{G^{i-1}}^2.$$

Assuming now that any element of \mathcal{T}_h^i has a bounded number of edges, independent of i , we arrive at

$$c_1 \|a\|_{G^i} \leq \sum_{K \in \mathcal{T}_h^i} \|a\|_{\Delta K} \leq c_2 \|a\|_{G^i},$$

where c_1, c_2 are positive constants. Using the result of [11], we also note that (2.9) gives

$$|\eta^{i+0}|_{\Omega^i}^2 \leq ch \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_{1,\Delta K} ch^0 \|\varepsilon\|_{1,\Delta K} + ch^{-1} h^2 \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_{1,\Delta K}^2 \leq ch \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_{1,\Delta K}^2. \quad (2.10)$$

We set

$$\mathcal{F} := c_0 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_{\Delta K}^2 + c_1 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\nabla \varepsilon\|_{\Delta K}^2 + \frac{1}{2} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} |\varepsilon^n|_{\Omega^n}^2 + \frac{1}{2} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} |u_h^{i+0} - u_h^i|_{\Omega^i}^2.$$

We observe that, for $0 < \epsilon \leq 1$, (2.6) and Young's inequality give

$$\int_\ell |\eta|^{1+\epsilon} ds \leq \int_\ell [c + |\eta|^2] ds \leq ch + \int_\ell |\eta|^2 ds \leq ch + ch \|\varepsilon\|_{1,\Delta K}^2,$$

which yields

$$\left(\int_\ell |\eta|^{1+\epsilon} ds \right)^{1/(1+\epsilon)} \leq ch^{1/(1+\epsilon)} + ch^{1/(1+\epsilon)} \|\varepsilon\|_{1,\Delta K}^{2/(1+\epsilon)}. \quad (2.11)$$

The same argument, applied on K and using the fact that the volume of K is of order $\mathcal{O}(h^2)$, yields

$$\|\eta\|_{L^{1+\epsilon}(K)} \leq ch^{2/(1+\epsilon)} + ch^{2/(1+\epsilon)} \|\varepsilon\|_{1,\Delta K}^{2/(1+\epsilon)}. \quad (2.12)$$

By Young's inequality we also have, for any $\gamma > 0$,

$$\gamma h^{1/(1+\epsilon)} \|\varepsilon\|_{1,\Delta K}^{2/(1+\epsilon)} \leq \gamma^{(1+\epsilon)/\epsilon} h + h^{1-\epsilon} \|\varepsilon\|_{1,\Delta K}^2, \quad (2.13)$$

$$\gamma h^{2/(1+\epsilon)} \|\varepsilon\|_{1,\Delta K}^{2/(1+\epsilon)} \leq \gamma^{(1+\epsilon)/\epsilon} h^2 + h^{2-2\epsilon} \|\varepsilon\|_{1,\Delta K}^2. \quad (2.14)$$

$$+ ch^{1/(1+\epsilon)} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \left(\int_{\ell} |[\nabla u_h \cdot \mathbf{n}]_{\ell}|^{\vartheta} ds \right)^{1/\vartheta},$$

for a positive constant c_0 , as small as desired and h sufficiently small. The result now follows from (2.5). \square

We assume now that the non-cylindrical domain S_T is two dimensional in the spatial variables and $\Omega(t)$ is of the form

$$\Omega(t) := \{(z, \theta) \in \mathbb{R}^2 : z \in (0, s(t, \theta)), \theta \in (\theta_1, \theta_2)\}, \quad \text{for } t \in [0, T], \quad (2.15)$$

where s is a smooth, positive function. Specific smoothness assumptions are given later. In this case, for g a sufficiently smooth function it follows that

$$\int_0^t \int_{\Omega(t)} g \, dx dt = \int_0^t \int_{\theta_1}^{\theta_2} \int_0^{s(t, \theta)} g \, dz d\theta dt.$$

For a given t , the boundary $\Gamma(t)$ of $\Omega(t)$ consists of the surfaces $z = 0$, $z = s(t, \theta)$, $\theta = \theta_1, \theta_2$. We apply the transformation [1, 4, 5, 7]

$$y := \frac{z}{s}, \quad \hat{u}(y, \theta, t) := u(z, \theta, t).$$

This change of variables gives $y \in (0, 1)$, $\Omega(t) \hookrightarrow \hat{\Omega} =: \hat{\Omega}^i$, for any i and $t \in (0, T)$, and $S_T \hookrightarrow \hat{S}_T$, $G^i \hookrightarrow \hat{G}^i$, where

$$\begin{aligned} \hat{\Omega} &= (0, 1) \times (\theta_1, \theta_2) \quad (t - \text{independent}), \\ \hat{S}_T &= (0, T) \times (0, 1) \times (\theta_1, \theta_2) \quad (\text{cylindrical}), \\ \hat{G}^i &= (t^i, t^{i+1}) \times \hat{\Omega}. \end{aligned}$$

The initial and boundary value problem (1.2) now becomes

$$\begin{aligned} \partial_t \hat{u} &= A \hat{u}_{yy} + B \hat{u}_{y\theta} + C \hat{u}_{\theta\theta} + D \hat{u}_y + \beta_1 \hat{u} + \hat{f} \quad \text{in } \hat{S}_T, \\ \hat{u} &= 0 \quad \text{at } y = 0, 1, \\ \hat{u} &= 0 \quad \text{at } \theta = \theta_1, \theta_2, \\ \hat{u}(y, \theta, 0) &= u^0(ys(0, \theta), \theta), \end{aligned} \quad (2.16)$$

where $\hat{u}(y, \theta, t) := u(z, \theta, t)$, $A = \frac{1}{s^2} + \frac{y^2 s_{\theta}^2}{s^2}$, $B = -2y \frac{s_{\theta}}{s}$, $C = 1$, $D = \frac{ys_t}{s} - \frac{y}{s^2} (s_{\theta\theta}s - 2s_{\theta}^2)$, $\beta_1 = \beta$ and $\hat{f}(y, \theta, t) = f(z, \theta, t)$. For the purpose of the analysis below we assume that

$$s, s_t, s_{\theta}, s_{\theta\theta} \in L^{\infty}((0, T) \times (\theta_1, \theta_2)), \quad (2.17)$$

which implies that the space-time coefficients of problem (2.16) satisfy $A, B, D \in L^{\infty}(\hat{S}_T)$. Also, C, β_1 are constants and thus in $L^{\infty}(\hat{S}_T)$. The next lemma presents an estimate for $\|\partial_t \hat{u}\|$.

Lemma 2.6. *Let \hat{u} be the solution of (2.16), and assume that (2.17) holds. Then, for any $t \leq T$, it holds*

$$\int_0^t \|\hat{u}\|_{H^2(\hat{\Omega})}^2 + \int_0^t \|\hat{u}_t\|^2 dt \leq c \|\hat{u}(0)\|_{H^1(\hat{\Omega})}^2 + c \int_0^t \|\hat{f}\|^2 dt, \quad (2.18)$$

provided that $u^0 \in H^1(\Omega(0))$ and $f \in L^2(G(0, T))$.

Proof. The regularity of s, s_{θ} yields $\hat{u}^0 \in H^{2m+1}(\hat{\Omega})$ for $m = 0$, provided that $u^0 \in H^1(\Omega(0))$. If $\hat{f} \in L^2(0, T; L^2(\hat{\Omega}))$, or, equivalently, since $s \in L^{\infty}((0, T) \times (\theta_1, \theta_2))$, that $f \in L^2(G(0, T))$, then relation (55), p. 365 of Evans [19], yields the result. \square

Since

$$\partial_t u = \partial_t \hat{u} - \frac{y}{s} s_t \hat{u}_y, \quad \hat{u}_y = s u_z, \quad \hat{u}_\theta = y s_\theta u_z + u_\theta,$$

we readily obtain from (2.18) the proof of the next proposition.

Proposition 2.7. *If u is the solution of (1.2) then, under the assumptions of Lemma 2.6,*

$$\sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\partial_t u\|_K^2 \leq c|u^0|_{\Omega^0}^2 + c|\nabla u^0|_{\Omega^0}^2 + c \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|f\|_K^2, \quad (2.19)$$

for any $1 \leq n \leq N$.

Remark 2.8. We note that the solution u has been assumed to be continuous in time in $(0, T)$. For this, it is sufficient that u belongs to $H^2(S_T)$, since $u = u(z, \theta, t)$ and $(z, \theta) \in \mathbb{R}^2$, $t \in \mathbb{R}$. For the former, it is sufficient that

$$\int_0^T [\|\hat{u}_{tt}\|_{L^2(\hat{\Omega})}^2 + \|\hat{u}_t\|_{H^1(\hat{\Omega})}^2 + \|\hat{u}\|_{H^2(\hat{\Omega})}^2] dt < \infty.$$

Observe that $\int_0^T \|\hat{u}\|_{H^2(\hat{\Omega})}^2 dt < \infty$, under the regularity assumptions on u^0 , f , s , in the proof of Lemma 2.6. The term $\|\hat{u}_t\|_{H^1(\hat{\Omega})}^2$ may be estimated by differentiating (2.16) in space, while $\|\hat{u}_{tt}\|_{L^2(\hat{\Omega})}^2$ may be estimated by differentiating (2.16) in time. Thus, $\int_0^T \|\hat{u}_t\|_{H^1(\hat{\Omega})}^2$ is bounded if $\hat{u} \in L^2(0, T; H^3(\hat{\Omega}))$, while $\|\hat{u}_{tt}\|_{L^2(\hat{\Omega})}^2$ is bounded if $\hat{u}^0 \in H^3(\hat{\Omega})$ and $\hat{f}_t \in L^2(0, T; L^2(\hat{\Omega}))$. The regularity of s , s_θ , $s_{\theta\theta}$ yields $\hat{u}^0 \in H^2(\hat{\Omega})$, if $u^0 \in H^2(\Omega(0))$. If we assume additionally, that $s_{tt}, s_{t\theta}, s_{t\theta\theta}, s_{\theta\theta\theta} \in L^\infty((0, T) \times (\theta_1, \theta_2))$, and that $\hat{f} \in L^2(0, T; H^1(\hat{\Omega}))$, $\hat{f}_t \in L^2(0, T; L^2(\hat{\Omega}))$, or, in the initial coordinates that $f, f_t, f_z, f_\theta \in L^2(G(0, T))$ then, we have $\hat{u} \in L^2(0, T; H^3(\hat{\Omega}))$.

In view of the remarks above, we conclude that if

$$u^0 \in H^3(\Omega(0)), \quad s, s_t, s_{tt}, s_{t\theta}, s_\theta, s_{\theta\theta}, s_{t\theta\theta}, s_{\theta\theta\theta} \in L^\infty(\hat{S}_T), \quad f, f_t, f_z, f_\theta \in L^2(G(0, T)), \quad (2.20)$$

then u is continuous, and obviously, also (2.19) is valid, i.e. $u_t \in L^2(G(0, T))$ and moreover, by (2.18), $u \in H^1(G(0, T))$. We write $u = u - u_h + u_h = \varepsilon + u_h$, $\partial_t \varepsilon = \partial_t u - \partial_t u_h$. Using (2.19), (2.20) and this splitting we obtain the following estimate for ε_t , where in place of T we may take any nodal point $t^n \leq T$.

Theorem 2.9. *Let u be the solution of (1.2) and $1 \leq n \leq N$. If*

$$u^0 \in H^3(\Omega(0)), \quad s, s_{tt}, s_t, s_{t\theta}, s_\theta, s_{\theta\theta}, s_{t\theta\theta}, s_{\theta\theta\theta} \in L^\infty(\hat{S}_T), \quad f, f_t, f_z, f_\theta \in L^2(G(0, t^n)), \quad (2.21)$$

then there exists a positive constant c_e such that

$$\sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\varepsilon_t\|_K^2 \leq c_e |u^0|_{\Omega^0}^2 + c_e |\nabla u^0|_{\Omega^0}^2 + c_e \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} (\|\partial_t u_h\|_K^2 + \|f\|_K^2), \quad (2.22)$$

and the solution u of the initial problem is continuous.

Applying (2.22) in (2.7), we have the following *a posteriori* upper bound.

Theorem 2.10. *Let $m = 2$ and $\Omega(t)$ be given by (2.15). Let u be the solution of (1.2) and $1 \leq n \leq N$. If (2.20) holds and $f \in L^\vartheta(G(0, t^n))$ then, for any $0 < \epsilon \leq 1$, and with $\vartheta := (1 + \epsilon)/\epsilon$, there exist positive constants*

c_1, c_2, c_3, c such that

$$\begin{aligned}
& \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\nabla \varepsilon\|_K^2 + c_1 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_K^2 + c_2 |\varepsilon^n|_{\Omega^n}^2 + c_3 \sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\Omega^i}^2 \\
& \leq c |\varepsilon^0|_{\Omega^0}^2 + ch^{2/(1+\epsilon)} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|f - \partial_t u_h + \Delta u_h + \beta u_h\|_{L^\vartheta(K)} \\
& \quad + ch^2 \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \|f - \partial_t u_h + \Delta u_h + \beta u_h\|_{L^\vartheta(K)}^\vartheta \\
& \quad + ch^{1/(1+\epsilon)} \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \left(\int_\ell |[\nabla u_h \cdot \mathbf{n}]_\ell|^\vartheta \, ds \right)^{1/\vartheta} \\
& \quad + ch \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} \sum_{\ell \in E_K^i \text{ in}} \int_\ell |[\nabla u_h \cdot \mathbf{n}]_\ell|^\vartheta \, ds \\
& \quad + ch^{1-\epsilon} \left[|u^0|_{\Omega^0}^2 + |\nabla u^0|_{\Omega^0}^2 + \sum_{i=0}^{n-1} \sum_{K \in \mathcal{T}_h^i} (\|\partial_t u_h\|_K^2 + \|f\|_K^2) \right],
\end{aligned} \tag{2.23}$$

for h sufficiently small.

Remark 2.11. Obviously, in view of the terms involving $h^{1-\epsilon}$ and $h^{1/(1+\epsilon)}$, the upper bound estimate (2.23) is suboptimal, viewed as an L^2 estimate. Moreover, (2.23) may not be seen as an H^1 estimate in the non-cylindrical domain, since the time derivative of the error (of the relevant H^1 norm) does not appear in the left-hand side. This is due to the fact that the space-time discretization is applied in the initial non-cylindrical domain, and therefore, the term ε_t reduces the order of convergence. As we shall demonstrate in the sequel, by transforming the initial problem into an equivalent one posed on a cylindrical domain, the *a posteriori* error analysis for the resulting scheme does indeed give optimal results in H^1 norm, see Theorem 3.6 and the remarks following that theorem.

3. A POSTERIORI ERROR ANALYSIS FOR THE EQUIVALENT PROBLEM

3.1. Change of variables – the general parabolic problem

As we showed in the analysis in Section 2, the initial problem (1.2) for a case of interest when $m = 2$, is transformed to an equivalent problem, *cf.* (2.16), posed on a cylindrical domain.

We shall consider the more general parabolic problem

$$\begin{aligned}
& \partial_t \hat{u} = A \hat{u}_{yy} + B \hat{u}_{y\theta} + C \hat{u}_{\theta\theta} + D \hat{u}_y + E \hat{u}_\theta + \beta_1 \hat{u} + \hat{f} \quad \text{in } \hat{S}_T, \\
& \hat{u} = 0 \quad \text{at } y = 0, 1, \\
& \hat{u} = 0 \quad \text{at } \theta = \theta_1, \theta_2, \\
& \hat{u}(y, \theta, 0) = \hat{u}^0(y, \theta),
\end{aligned} \tag{3.1}$$

with A, B, C, D, E, β_1 functions of (y, θ, t) , and A, C positive while $B^2 < 4AC$. Note that (2.16) is a special case of this general parabolic problem, with $E = 0$. Also, we shall assume that β_1 is a negative function whose absolute value is controlled by the transformation $\hat{u} \rightarrow e^{ct} \hat{u}$, for a suitable constant c . In the sequel we shall assume that the coefficients A, B, C, D, E, β_1 , and \hat{u}^0, \hat{f} , satisfy, for $1 \leq n \leq N$,

$$\begin{aligned}
& \hat{u}^0 \in H^3(\hat{\Omega}), \quad \hat{f}, \hat{f}_t, \hat{f}_y, \hat{f}_\theta \in L^2(\hat{S}_{t^n}), \\
& \partial_t^k \partial_x^\mu \mathcal{L} \in L^\infty(\hat{S}_{t^n}) \quad \text{for } \mathcal{L} := A, B, C, D, E, \beta_1, \quad x = y, \theta, \quad \text{and } k, \mu = 0, 1,
\end{aligned} \tag{3.2}$$

in $\hat{S}_{t^n} = (0, t^n) \times (0, 1) \times (\theta_1, \theta_2)$. These assumptions yield that u is continuous in time in $[0, t^n]$, and $u_t \in L^2(\hat{S}_{t^n})$, $u \in H^1(\hat{S}_{t^n})$, as in Remark 2.8. For the problem (3.1) we define an analogous to (1.3) space-time, discontinuous in time scheme, and seek $u_h \in V_h$ satisfying

$$\begin{aligned} \hat{B}_n(u_h, v_h) &= ((\hat{f}, v_h))_{\hat{G}^n}, \quad \forall v_h \in V_h^n, \quad n = 0, \dots, N-1, \\ u_h^0 &= \hat{u}^0, \end{aligned} \quad (3.3)$$

where, now, the bilinear form $\hat{B}_n(u_h, v_h)$ is defined as

$$\begin{aligned} \hat{B}_n(u_h, v_h) &:= -((u_h, \partial_t v_h))_{\hat{G}^n} - \int_{t^n}^{t^{n+1}} \mathcal{B}(u_h, v_h) dt \\ &\quad + (u_h^{n+1}, v_h^{n+1})_{\hat{\Omega}} - (u_h^n, v_h^{n+0})_{\hat{\Omega}} - ((\tilde{D}u_{hy} + \tilde{E}u_{h\theta} + \beta_1 u_h, v_h))_{\hat{G}^n}. \end{aligned} \quad (3.4)$$

Here, $\tilde{D} = D - A_y - \frac{B_\theta}{2}$ and $\tilde{E} = E - C_\theta - \frac{B_y}{2}$, while $\mathcal{B} : H^1(\hat{\Omega}) \times H^1(\hat{\Omega}) \rightarrow R$ is a bilinear form given by

$$\mathcal{B}(v, w) := -(Av_y, w_y)_{\hat{\Omega}} - (Cv_\theta, w_\theta)_{\hat{\Omega}} - \frac{1}{2} \left\{ (Bv_y, w_\theta)_{\hat{\Omega}} + (Bv_\theta, w_y)_{\hat{\Omega}} \right\}.$$

We first observe that, since $A, C > 0$ and $B^2 < 4AC$, there exists a positive constant c such that

$$-\mathcal{B}(u, u) \geq c|\nabla u|_{\hat{\Omega}}^2.$$

The proof of the following existence theorem is now immediate.

Theorem 3.1. *The problem (3.3) admits a unique solution $u_h \in V_h$.*

Proof. We set $v_h = u_h$ in (3.3) so that

$$-\int_{t^n}^{t^{n+1}} \mathcal{B}(u_h, u_h) dt - ((\tilde{D}u_{hy} + \tilde{E}u_{h\theta} + \beta_1 u_h, u_h))_{\hat{G}^n} + \frac{1}{2}|u_h^{n+1}|_{\hat{\Omega}}^2 \leq \frac{1}{2}|u_h^n|_{\hat{\Omega}}^2 + |((\hat{f}, u_h))_{\hat{G}^n}|. \quad (3.5)$$

Then, (3.5) gives by summation

$$\|u_h\|_{\hat{G}(0, t^n)}^2 + \|\nabla u_h\|_{\hat{G}(0, t^n)}^2 \leq c|\hat{u}^0|_{\hat{\Omega}}^2 + c\|\hat{f}\|_{\hat{G}(0, t^n)}^2,$$

for β_1 sufficiently negative. This establishes the existence of a unique solution of problem (3.3). \square

Let u_h be the unique solution of (3.3) in V_h and $\varepsilon := \hat{u} - u_h$, where \hat{u} is the solution of the continuous problem (3.1). Also, let $\eta := \varepsilon - v_h$, for v_h an arbitrary element of V_h . By computations entirely analogous to those leading to (2.4) we obtain

$$\begin{aligned} &-\sum_{i=0}^{n-1} ((\tilde{D}\varepsilon_y + \tilde{E}\varepsilon_\theta + \beta_1 \varepsilon, \varepsilon))_{\hat{G}^i} - \int_0^{t^n} \mathcal{B}(\varepsilon, \varepsilon) dt + \frac{1}{2}|\varepsilon^n|_{\hat{\Omega}}^2 + \frac{1}{2} \sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\hat{\Omega}}^2 \\ &= \frac{1}{2}|\varepsilon^0|_{\hat{\Omega}}^2 + ((\hat{f}, \eta))_{\hat{G}(0, t^n)} + \sum_{i=0}^{n-1} ((u_h, \partial_t \eta))_{\hat{G}^i} \\ &\quad + \sum_{i=0}^{n-1} \int_{t^i}^{t^{i+1}} \mathcal{B}(u_h, \eta) dt + (u_h^0, \eta^0)_{\hat{\Omega}} - (u_h^n, \eta^n)_{\hat{\Omega}} \\ &\quad - \sum_{i=0}^{n-1} (u_h^i, \eta^i - \eta^{i+0})_{\hat{\Omega}} + \sum_{i=0}^{n-1} ((\tilde{D}u_{hy} + \tilde{E}u_{h\theta} + \beta_1 u_h, \eta))_{\hat{G}^i}. \end{aligned} \quad (3.6)$$

Given the partition $0 = t^0 < t^1 \cdots < t^N = T$ of $[0, T]$, we let $I_n = (t^n, t^{n+1}]$, $k_n = t^{n+1} - t^n$ and $k_n = \max_{0 \leq i \leq n-1} k_i$. Let us consider a triangulation T_h^n of $\hat{\Omega}$ and a finite element space S_h^n in each interval I_n consisting of functions in $H_0^1(\hat{\Omega})$ which reduce to polynomials of degree $r-1$ on each triangle. Here, for any n , h_n is the maximum diameter of elements of the triangulation T_h^n . As before, we let $h_n = \max_{0 \leq i \leq n-1} h_i$. Now let $\mathcal{V}_{hk} = \mathcal{V}_{hk}(q)$, $q \geq 2$, denote the space of piecewise polynomial functions $\phi : (0, T] \times \hat{\Omega} \rightarrow \mathbb{R}$ of the form: $\phi|_{I_n \times \hat{\Omega}} = \sum_{j=0}^{q-1} t^j \chi_j(y)$, $\chi_j \in S_h^n$. It follows that the functions of \mathcal{V}_{hk} are, for each $t \in I_n$ elements of S_h^n and for each $y \in \hat{\Omega}$ piecewise polynomial functions of degree $q-1$, with possible discontinuities at the temporal nodes t^n , $n = 0, \dots, N-1$. We also define $\mathcal{V}_{hk}^n = \{\phi|_{I_n \times \hat{\Omega}} : \phi \in \mathcal{V}_{hk}\}$.

3.2. Upper bound

Let us define the symmetric matrix

$$\mathcal{M} := \begin{pmatrix} A & \frac{B}{2} \\ \frac{B}{2} & C \end{pmatrix}.$$

If ℓ is an interior edge of T_h^n , let $[\mathcal{M}\nabla u_h \cdot \mathbf{n}]_\ell$ denote the jump of $\mathcal{M}\nabla u_h \cdot \mathbf{n}$ across the edge ℓ , where \mathbf{n} is the normal direction, *i.e.*

$$[\mathcal{M}\nabla u_h \cdot \mathbf{n}]_\ell := \mathcal{M}\nabla u_h \cdot \mathbf{n}|_{\ell^+} - \mathcal{M}\nabla u_h \cdot \mathbf{n}|_{\ell^-}.$$

Theorem 3.2. *For the error $\varepsilon = \hat{u} - u_h$ it holds*

$$\begin{aligned} & - \sum_{i=0}^{n-1} ((\tilde{D}\varepsilon_y + \tilde{E}\varepsilon_\theta + \beta_1\varepsilon, \varepsilon))_{\hat{G}^i} - \int_0^{t^n} \mathcal{B}(\varepsilon, \varepsilon) dt + \frac{1}{2} |\varepsilon^n|_{\hat{\Omega}}^2 + \frac{1}{2} \sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\hat{\Omega}}^2 \\ &= \frac{1}{2} |\varepsilon^0|_{\hat{\Omega}}^2 + \sum_{i=0}^{n-1} \int_{I_i} \sum_{K \in T_h^i} (\hat{f} - \partial_t u_h + Au_{hyy} + Bu_{hy\theta} \\ & \quad + Cu_{h\theta\theta} + Du_{hy} + Eu_{h\theta} + \beta_1 u_h, \eta)_K dt \\ & \quad - \sum_{i=0}^{n-1} \int_{I_i} \sum_{K \in T_h^i} \sum_{\ell \in E_{K_{\text{in}}}^i} \int_{\ell} \eta [\mathcal{M}\nabla u_h \cdot \mathbf{n}]_\ell ds dt + \sum_{i=0}^{n-1} (u_h^i - u_h^{i+0}, \eta^{i+0})_{\hat{\Omega}}, \end{aligned} \quad (3.7)$$

where $(\cdot, \cdot)_K$ is the usual inner product in $L^2(K)$.

For the purpose of establishing an error bound, we shall take $r = 2$, so that S_h^n consists of piecewise linear functions in the y, θ variables. We also assume that T_h^n is regular, uniformly on any n . For a fixed n and $t \in I_n$, we denote by $\pi_h^n \varepsilon(t)$ the Clément interpolant of the error $\varepsilon(t)$ in $P_1(K)$, where K is an element of T_h^n . Since $r = 2$, we have $\pi_h^n \varepsilon(t) \in S_h^n$. In fact, if $B_n = \{b_i\}_{i=1}^{\dim S_h^n}$ is a basis of S_h^n , for any $y \in \hat{\Omega}$,

$$\pi_h^n \varepsilon(y, \theta, t) = \sum_{i=1}^{\dim S_h^n} c_i(t) b_i(y, \theta),$$

where $c_i(t) \in \mathbb{R}$. In order to obtain an element of \mathcal{V}_{hk} we define $v_h|_{\hat{G}^n}$ as the $L^2(I_n)$ -projection of $\pi_h^n \varepsilon$ in $P_{q-1}(I_n)$, the polynomials of order at most $q-1$ in $t \in I_n$. We shall also make use of the L^2 projection operator $L^n : L^2(I_n) \rightarrow P_{q-1}(I_n)$ defined by

$$\int_{I_n} t^j L^n(g(t)) dt = \int_{I_n} t^j g(t) dt, \quad \forall g \in L^2(I_n), \quad j = 0, \dots, q-1. \quad (3.8)$$

Let $f^{(q)}$ denote the q -time derivative of f . If $f(\cdot, t) \in C^q(\bar{I}_n)$ then, cf. [4],

$$\int_{I_n} |L^n f - f|^2 dt \leq c k_n^{2q+1} \left(\max_{t \in \bar{I}_n} |f^{(q)}(\cdot, t)| \right)^2. \quad (3.9)$$

We can now define $v_h \in \mathcal{V}_{hk}$ by

$$v_h|_{\mathcal{V}_{hk}^n} = L^n \circ \pi_h^n \varepsilon \text{ in } I_n \times \hat{\Omega}, \quad n = 0, \dots, N-1. \quad (3.10)$$

Obviously, it holds that

$$L^n \circ \pi_h^n \varepsilon(y, \theta, t) = \sum_{i=1}^{\dim S_h^n} L^n(c_i(t)) b_i(y, \theta).$$

As before, we note that there exists a positive constant C such that

$$\begin{aligned} \|\varepsilon - \pi_h^n \varepsilon\|_{L^2(K)} &\leq C \mathbf{h}_n \|\varepsilon\|_{H^1(\Delta K)}, \\ \|\varepsilon - \pi_h^n \varepsilon\|_{L^2(\ell)} &\leq C \mathbf{h}_n^{1/2} \|\varepsilon\|_{H^1(\Delta K)}. \end{aligned} \quad (3.11)$$

We define the $L^2(0, s; H^k(\hat{\Omega}))$ norm as

$$\|g\|_{L^2(0, s; H^k(\hat{\Omega}))} := \left(\int_0^s \|g(\cdot, t)\|_{H^k(\hat{\Omega})}^2 dt \right)^{1/2}, \quad k = 0, 1, \dots$$

where $H^0(\hat{\Omega}) := L^2(\hat{\Omega})$.

Remark 3.3. Considering the solution \hat{u} of the problem (3.1), if (3.2) holds also for $k, \mu = 0, \dots, m+1$, for some $m \geq 0$, then, for $1 \leq n \leq N$, and any $0 < s \leq t^n$, we have

$$\sum_{k=0}^{m+1} \|\hat{u}^{(k)}\|_{L^2(0, s; H^{2m+2-2k}(\hat{\Omega}))} \leq c \left(\sum_{k=0}^m \|\hat{f}^{(k)}\|_{L^2(0, s; H^{2m-2k}(\hat{\Omega}))} + \|\hat{u}^0\|_{H^{2m+1}(\hat{\Omega})} \right), \quad (3.12)$$

if $\hat{u}^0 \in H^{2m+1}(\hat{\Omega})$, and $\hat{f}^{(k)} \in L^2(0, s; H^{2m-2k}(\hat{\Omega}))$ for any $k = 0, \dots, m+1$, where we have set $\hat{u}^{(k)} := \partial_t^k \hat{u}$, $f^{(k)} := \partial_t^k f$, see, e.g. [19]. We shall make use of these estimates in the sequel. The error η of the interpolant $L^n \circ \pi_h^n \varepsilon$ of $\varepsilon = \hat{u} - u_h$ will result in high order space-time derivatives of the error ε . The restriction of $\partial_t^q \partial_x^\mu \varepsilon$ to an element of the partition may be written as $\partial_t^q \partial_x^\mu \hat{u} - \partial_t^q \partial_x^\mu u_h$. We shall estimate the derivatives $\partial_t^q \partial_x^\mu \hat{u}$ by (3.12), while $\partial_t^q \partial_x^\mu u_h$ will be used in the estimator.

In view of (3.7), we shall need to control the terms $|\eta^{i+0}|_{\hat{\Omega}}$, $\|\eta\|_{L^2(K)}$, and $\|\eta\|_{L^2(\ell)}$, so as to establish an *a posteriori* estimate. This is accomplished in the next two lemmata assuming a regular partition in t , i.e. that $\frac{k_i}{k_j} \geq c$ for any i, j , where c is some positive constant.

Lemma 3.4. *For any $1 \leq n \leq N$, it holds that*

$$\begin{aligned}
\sum_{i=0}^{n-1} |\eta^{i+0}|_{\hat{\Omega}}^2 &\leq c \mathbf{h}_n^2 \mathbf{k}_n^{-1} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\varepsilon\|_{H^1(K)}^2 dt + c \mathbf{h}_n^2 \mathbf{k}_n \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_{ht}\|_{H^1(K)}^2 dt \\
&+ c \mathbf{k}_n^{2q-1} \mathbf{h}_n^2 \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{H^1(K)}^2 dt \right] \\
&+ c \mathbf{k}_n^{2q+1} \mathbf{h}_n^2 \left[\|\hat{u}^0\|_{H^{2q+3}(\hat{\Omega})}^2 + \sum_{k=0}^{q+1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q+2-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{H^1(K)}^2 dt \right] \\
&+ c \mathbf{k}_n^{2q-1} \left[\|\hat{u}^0\|_{H^{2q-1}(\hat{\Omega})}^2 + \sum_{k=0}^{q-1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 dt \right] \\
&+ c \mathbf{k}_n^{2q+1} \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt \right] \\
&+ c \mathbf{h}_n^2 \mathbf{k}_n \left[\|\hat{u}^0\|_{H^3(\hat{\Omega})}^2 + \sum_{k=0}^1 \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2-2k}(\hat{\Omega}))}^2 \right]. \tag{3.13}
\end{aligned}$$

Proof. Observe first that for g smooth in t it follows that [4]

$$\max_{t \in \overline{I}_i} |L^i(g) - g| \leq c k_i^q \max_{t \in \overline{I}_i} |g^{(q)}(\cdot, t)|.$$

Furthermore, if v is smooth then $(\pi_h^i v)^{(q)} = \pi_h^i(v^{(q)})$. Indeed, $\dim \hat{\Omega} = 2$, and the domain is rectangular, so π_h^i is the standard Clemént interpolant defined via the local L^2 projections in space. Thus, the operator π_h^i is independent of t for any i , since $\hat{\Omega}$ is time independent. We write

$$\eta^{i+0} = \varepsilon(\cdot, t^{i+0}) - L^i \circ \pi_h^i \varepsilon(\cdot, t^{i+0}) = \varepsilon(\cdot, t^{i+0}) - \pi_h^i \varepsilon(\cdot, t^{i+0}) - (L^i - I) \pi_h^i \varepsilon(\cdot, t^{i+0}),$$

and use the previous estimates together with (3.11) and the local Sobolev inequality (A.3) for $t \in I_i$, to obtain

$$|\eta^{i+0}|_{\hat{\Omega}}^2 \leq c \mathbf{h}_n^2 \sum_{K \in T_h^i} \int_{I_i} \left[k_i^{-1} \|\varepsilon\|_{H^1(\Delta K)}^2 + k_i \|\varepsilon_t\|_{H^1(\Delta K)}^2 \right] dt + c k_i^{2q} \int_0^1 \int_{\theta_1}^{\theta_2} (\max_{t \in \overline{I}_i} |\pi_h^i(\varepsilon^{(q)})(\cdot, t)|)^2 d\theta dy.$$

Using again (A.3), we get

$$\begin{aligned}
|\pi_h^i(\varepsilon^{(q)})(t)| &\leq |\pi_h^i(\varepsilon^{(q)})(t) - \varepsilon^{(q)}(t)| + |\varepsilon^{(q)}(t)| \leq c \left(k_i^{-1} \int_{I_i} [|\pi_h^i(\varepsilon^{(q)}) - \varepsilon^{(q)}|^2 + |\varepsilon^{(q)}|^2] dt \right)^{1/2} \\
&+ c \left(k_i \int_{I_i} [|\pi_h^i(\varepsilon^{(q+1)}) - \varepsilon^{(q+1)}|^2 + |\varepsilon^{(q+1)}|^2] dt \right)^{1/2}.
\end{aligned}$$

Splitting the time derivatives of $\varepsilon = \hat{u} - u_h$ and using the above inequalities, we obtain

$$\begin{aligned} |\eta^{i+0}|_{\hat{\Omega}}^2 &\leq c \mathbf{h}_n^2 k_i^{-1} \sum_{K \in T_h^i} \int_{I_i} \|\varepsilon\|_{H^1(K)}^2 dt + c \mathbf{h}_n^2 k_i \sum_{K \in T_h^i} \int_{I_i} \|u_{ht}\|_{H^1(K)}^2 dt \\ &\quad + c k_i^{2q-1} \mathbf{h}_n^2 \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{H^1(K)}^2 dt + c k_i^{2q+1} \mathbf{h}_n^2 \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{H^1(K)}^2 dt \\ &\quad + c k_i^{2q-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 dt + c k_i^{2q+1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt + \mathcal{A}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{A} &\leq c \mathbf{h}_n^2 k_i \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}_t\|_{H^1(K)}^2 dt + c k_i^{2q-1} \mathbf{h}_n^2 \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q)}\|_{H^1(K)}^2 dt + c k_i^{2q+1} \mathbf{h}_n^2 \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q+1)}\|_{H^1(K)}^2 dt \\ &\quad + c k_i^{2q-1} \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q)}\|_{L^2(K)}^2 dt + c k_i^{2q+1} \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q+1)}\|_{L^2(K)}^2 dt. \end{aligned}$$

Summing over i and applying the estimates (3.12) for the time derivatives of \hat{u} we obtain (3.13). \square

Lemma 3.5. *For any $1 \leq n \leq N$, it holds that*

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\eta\|_{L^2(K)}^2 dt &\leq c \mathbf{h}_n^2 \left(\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\varepsilon\|_{L^2(K)}^2 dt + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\nabla \varepsilon\|_{L^2(K)}^2 dt \right) \\ &\quad + c \mathbf{k}_n^{2q} \left[\|\hat{u}^0\|_{H^{2q-1}(\hat{\Omega})}^2 \sum_{k=0}^{q-1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2-2k}(\hat{\Omega}))}^2 \right] \\ &\quad + c \mathbf{k}_n^{2q+2} \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2 \right] \\ &\quad + c \mathbf{k}_n^{2q} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 dt \\ &\quad + c \mathbf{k}_n^{2q+2} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt, \end{aligned} \tag{3.14}$$

and

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|\eta\|_{L^2(\ell)}^2 dt &\leq c \mathbf{h}_n \left(\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\varepsilon\|_{L^2(K)}^2 dt + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\nabla \varepsilon\|_{L^2(K)}^2 dt \right) \\ &\quad + c \mathbf{k}_n^{2q} \mathbf{h}_n^{-1} \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2 \right] \\ &\quad + c \mathbf{k}_n^{2q+2} \mathbf{h}_n^{-1} \left[c \|\hat{u}^0\|_{H^{2q+3}(\hat{\Omega})}^2 + c \sum_{k=0}^{q+1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q+2-2k}(\hat{\Omega}))}^2 \right] \\ &\quad + c \mathbf{k}_n^{2q} \mathbf{h}_n^{-1} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(\ell)}^2 dt \\ &\quad + c \mathbf{k}_n^{2q+2} \mathbf{h}_n^{-1} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(\ell)}^2 dt. \end{aligned} \tag{3.15}$$

Proof. By (3.9) and using the fact that L^i is the L^2 projection, we have for $\mathcal{A} = K, \ell$,

$$\begin{aligned} \int_{I_i} \|\eta\|_{L^2(\mathcal{A})}^2 dt &\leq 2 \int_{I_i} \|\varepsilon - L^i \varepsilon\|_{L^2(\mathcal{A})}^2 dt + 2 \int_{I_i} \|L^i \varepsilon - L^i \circ \pi_h^i \varepsilon\|_{L^2(\mathcal{A})}^2 dt \\ &\leq 2k_i^{2q+1} \max_{t \in \bar{I}_i} |\varepsilon^{(q)}(\cdot, t)|^2_{L^2(\mathcal{A})} + 2 \int_{I_i} \|\varepsilon - \pi_h^i \varepsilon\|_{L^2(\mathcal{A})}^2 dt \\ &\leq 2k_i^{2q+1} \max_{t \in \bar{I}_i} |\varepsilon^{(q)}(\cdot, t)|^2_{L^2(\mathcal{A})} + ch_{\mathcal{A}}^2 \left(2 \int_{I_i} \|\varepsilon\|_{L^2(\Delta K)}^2 dt + 2 \int_{I_i} \|\nabla \varepsilon\|_{L^2(\Delta K)}^2 dt \right), \end{aligned} \quad (3.16)$$

where $h_{\mathcal{A}} = h_n$ if $\mathcal{A} = K$, or $h_{\mathcal{A}} = h_n^{1/2}$ if $\mathcal{A} = \ell$. Note that by a local Sobolev inequality in \mathbb{R} (cf. (A.3)) it holds that

$$\max_{t \in \bar{I}_i} |\varepsilon^{(q)}(\cdot, t)|^2 \leq ck_i^{-1} \int_{I_i} |\varepsilon^{(q)}(\cdot, t)|^2 dt + ck_i \int_{I_i} |\varepsilon^{(q+1)}(\cdot, t)|^2 dt,$$

and thus

$$\begin{aligned} \|\max_{t \in \bar{I}_i} |\varepsilon^{(q)}(\cdot, t)|\|_{L^2(\mathcal{A})}^2 &\leq ck_i^{-1} \int_{I_i} \|\hat{u}^{(q)}\|_{L^2(\mathcal{A})}^2 dt + ck_i^{-1} \int_{I_i} \|u_h^{(q)}\|_{L^2(\mathcal{A})}^2 dt \\ &\quad + ck_i \int_{I_i} \|\hat{u}^{(q+1)}\|_{L^2(\mathcal{A})}^2 dt + ck_i \int_{I_i} \|u_h^{(q+1)}\|_{L^2(\mathcal{A})}^2 dt. \end{aligned} \quad (3.17)$$

Substituting in (3.16) we obtain

$$\begin{aligned} \int_{I_i} \|\eta\|_{L^2(K)}^2 dt &\leq 2ck_i^{2q} \int_{I_i} \|\hat{u}^{(q)}\|_{L^2(K)}^2 dt + 2ck_i^{2q} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 dt \\ &\quad + 2ck_i^{2q+2} \int_{I_i} \|\hat{u}^{(q+1)}\|_{L^2(K)}^2 dt + 2ck_i^{2q+2} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt \\ &\quad + ch_n^2 \left(2 \int_{I_i} \|\varepsilon\|_{L^2(\Delta K)}^2 dt + 2 \int_{I_i} \|\nabla \varepsilon\|_{L^2(\Delta K)}^2 dt \right). \end{aligned} \quad (3.18)$$

Using the local trace inequality (A.1)

$$\|\hat{u}^{(q)}\|_{L^2(\ell)}^2 \leq \int_{\partial K} |\hat{u}^{(q)}|^2 ds \leq ch_n^{-1} \|\hat{u}^{(q)}\|_{H^1(K)}^2,$$

we obtain

$$\begin{aligned} \int_{I_i} \|\eta\|_{L^2(\ell)}^2 dt &\leq 2ck_i^{2q} h_n^{-1} \int_{I_i} \|\hat{u}^{(q)}\|_{H^1(K)}^2 dt + 2ck_i^{2q} h_n^{-1} \int_{I_i} \|u_h^{(q)}\|_{L^2(\ell)}^2 dt \\ &\quad + 2ck_i^{2q+2} h_n^{-1} \int_{I_i} \|\hat{u}^{(q+1)}\|_{H^1(K)}^2 dt + 2ck_i^{2q+2} h_n^{-1} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(\ell)}^2 dt \\ &\quad + ch_n \left(2 \int_{I_i} \|\varepsilon\|_{L^2(\Delta K)}^2 dt + 2 \int_{I_i} \|\nabla \varepsilon\|_{L^2(\Delta K)}^2 dt \right). \end{aligned} \quad (3.19)$$

Using (3.12) we arrive at

$$\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q)}\|_{L^2(K)}^2 \leq \|\hat{u}^{(q)}\|_{L^2(0, t^n; L^2(\hat{\Omega}))}^2 \leq c \|\hat{u}^0\|_{H^{2q-1}(\hat{\Omega})}^2 + c \sum_{k=0}^{q-1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2-2k}(\hat{\Omega}))}^2,$$

$$\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|\hat{u}^{(q+1)}\|_{L^2(K)}^2 \leq \|\hat{u}^{(q+1)}\|_{L^2(0, t^n; L^2(\hat{\Omega}))}^2 \leq c \|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + c \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2, \quad (3.20)$$

while

$$\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|\hat{u}^{(q)}\|_{H^1(K)}^2 \leq \|\hat{u}^{(q)}\|_{L^2(0, t^n; H^1(\hat{\Omega}))}^2 \leq c \|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + c \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2,$$

and

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|\hat{u}^{(q+1)}\|_{H^1(K)}^2 dt &\leq \|\hat{u}^{(q+1)}\|_{L^2(0, t^n; H^1(\hat{\Omega}))}^2 \\ &\leq c \|\hat{u}^0\|_{H^{2q+3}(\hat{\Omega})}^2 + c \sum_{k=0}^{q+1} \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q+2-2k}(\hat{\Omega}))}^2. \end{aligned} \quad (3.21)$$

Summing over i and K in (3.18) and summing over i, K, ℓ in (3.19), and by substituting in (3.20) and (3.21), respectively, we obtain the estimates (3.14) and (3.15). \square

The following theorem presents an upper bound for the *a posteriori* error. Its proof follows by using in (3.7) the fact that β_1 may be suitably chosen and the estimates (3.13)–(3.15).

Theorem 3.6. *If the solution \hat{u} of problem (3.1), satisfies (3.2) for $k, \mu = 0, \dots, m+1$, for $m = q$, and if $h_n^2 k_n^{-1}$ is sufficiently small, then the error $\varepsilon = \hat{u} - u_h$ satisfies*

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \left\{ \|\varepsilon\|_K^2 + \|\nabla \varepsilon\|_K^2 \right\} + |\varepsilon^n|_{\hat{\Omega}}^2 + \sum_{i=0}^{n-1} |u_h^{i+0} - u_h^i|_{\hat{\Omega}}^2 \\ \leq c |\varepsilon^0|_{\hat{\Omega}}^2 + \left[c h_n^2 \right] \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \|\hat{f} - \partial_t u_h + A u_h y y\|_{L^2(K)}^2 \\ + \left[c h_n \right] \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{\ell} \left| [\mathcal{M} \nabla u_h \cdot \mathbf{n}]_{\ell} \right|^2 ds \\ + \left[c k_n^{2q} h_n^{-2} \right] \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(\ell)}^2 \\ + \left[c k_n^{2q+2} h_n^{-2} \right] \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \sum_{\ell \in E_K^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(\ell)}^2 dt \\ + \left[c h_n^2 k_n \right] \mathcal{E}_1 + \left[c k_n^{2q-1} h_n^2 \right] \mathcal{E}_2 + \left[c k_n^{2q+1} h_n^2 \right] \mathcal{E}_3 + \left[c k_n^{2q-1} \right] \mathcal{E}_4 \\ + \left[c k_n^{2q+1} \right] \mathcal{E}_5 + \left[c k_n^{2q} h_n^{-2} \right] \mathcal{E}_6 + \left[c k_n^{2q+2} h_n^{-2} \right] \mathcal{E}_7, \end{aligned}$$

where $\mathcal{E}_i = \mathcal{E}_i(\hat{u}^0, \hat{f}, u_h)$, $i = 1, \dots, 7$, are given by

$$\mathcal{E}_1 := \left[\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_{ht}\|_{H^1(K)}^2 dt + \|\hat{u}^0\|_{H^3(\hat{\Omega})}^2 + \sum_{k=0}^1 \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2-2k}(\hat{\Omega}))}^2 \right],$$

$$\mathcal{E}_2 := \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0, t^n; H^{2q-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{H^1(K)}^2 dt \right],$$

$$\begin{aligned}
\mathcal{E}_3 &:= \left[\|\hat{u}^0\|_{H^{2q+3}(\hat{\Omega})}^2 + \sum_{k=0}^{q+1} \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q+2-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{H^1(K)}^2 dt \right], \\
\mathcal{E}_4 &:= \left[\|\hat{u}^0\|_{H^{2q-1}(\hat{\Omega})}^2 + \sum_{k=0}^{q-1} \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q-2-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 dt \right], \\
\mathcal{E}_5 &:= \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt \right], \\
\mathcal{E}_6 &:= \left[\|\hat{u}^0\|_{H^{2q-1}(\hat{\Omega})}^2 \sum_{k=0}^{q-1} \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q-2-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q)}\|_{L^2(K)}^2 \right. \\
&\quad \left. + \|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q-2k}(\hat{\Omega}))}^2 \right], \\
\mathcal{E}_7 &:= \left[\|\hat{u}^0\|_{H^{2q+1}(\hat{\Omega})}^2 + \sum_{k=0}^q \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q-2k}(\hat{\Omega}))}^2 + \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \int_{I_i} \|u_h^{(q+1)}\|_{L^2(K)}^2 dt \right. \\
&\quad \left. + \|\hat{u}^0\|_{H^{2q+3}(\hat{\Omega})}^2 + \sum_{k=0}^{q+1} \|\hat{f}^{(k)}\|_{L^2(0,t^n;H^{2q+2-2k}(\hat{\Omega}))}^2 \right],
\end{aligned}$$

for any $1 \leq n \leq N$. It is assumed that $\hat{f}^{(k)} \in L^2(0, s; H^{2m-2k}(\hat{\Omega}))$ for any $k = 0, \dots, m+1$, $m = q-1, q, q+1$, and that $\hat{u}^0 \in H^{2q+3}(\hat{\Omega})$.

Several remarks concerning the nature of this estimate are in order:

- (1) The assumption $\mathbf{h}_n^2 \mathbf{k}_n^{-1} \leq c$, when $\mathcal{O}(\mathbf{h}_n) = \mathcal{O}(\mathbf{k}_n)$, is transformed into a smallness assumption for the space discretization parameter \mathbf{h}_n .
- (2) The estimate of Theorem 3.6 involves the L^2 and H^1 norms of the error ε . Note that since $m = 2$, the lengths of the interior edges $|\ell|$ are of order $\mathcal{O}(\mathbf{h}_n)$ while the areas of the elements $|K|$ of the partition are of order $\mathcal{O}(\mathbf{h}_n^2)$. Thus, the *a posteriori* estimate of Theorem 3.6 is optimal in space, as an H^1 estimate, if the $\|\cdot\|_{L^2(K)}^2$, $\|\cdot\|_{H^1(K)}^2$ coefficients are of order at least $\mathcal{O}(\mathbf{h}_n^2)$, and if the $\|\cdot\|_{L^2(\ell)}^2$ and trace term coefficients involving \mathcal{M} are of order at least $\mathcal{O}(\mathbf{h}_n)$. So, if $\mathcal{O}(\mathbf{k}_n) = \mathcal{O}(\mathbf{h}_n)$, then $q \geq 2$ provides an optimal result in H^1 . However, if $\mathcal{O}(\mathbf{k}_n) = \mathcal{O}(\mathbf{h}_n)$, and $q \geq 2$, the estimate of Theorem 3.6 is sub-optimal in space as an L^2 estimate, and of total order \mathbf{h}_n , due to the trace terms. We have not estimated the constants appearing in the upper bound. However, attention must be given to the Gronwall constants that multiply the error terms \mathcal{E}_i that involve \hat{f} and \hat{u}^0 , for they may be large.
- (3) Assume that $\mathcal{O}(\mathbf{k}_n) = \mathcal{O}(\mathbf{h}_n)$. The aforementioned constants appear as coefficients of the terms involving $c\mathbf{h}_n^3, c\mathbf{h}_n^{2q+1}, c\mathbf{h}_n^{2q+3}, c\mathbf{h}_n^{2q-1}, c\mathbf{h}_n^{2q+1}, c\mathbf{h}_n^{2q-2}, c\mathbf{h}_n^{2q}$. If we select $q \geq 3$ then, the lowest order term is of order at least $\mathcal{O}(\mathbf{h}_n^3)$. Hence, for \mathbf{h}_n small, all these coefficients are less than $c\mathbf{h}_n^3 < 1 \cdot \mathbf{h}_n^2$. Thus, the error estimate remains optimal in space in H^1 and the resulting *a posteriori* constants are controlled by the constant one.
- (4) When $m \geq 3$, our analysis may be easily extended to space-time, discontinuous in time Galerkin methods applied to parabolic initial and boundary value problems of general type, with Dirichlet boundary conditions, such as

$$\begin{aligned}
\hat{u}_t &= \mathcal{L}(\hat{u}) + F \quad \text{in } (0, T) \times \tilde{\Omega}, \\
\hat{u}(\cdot, 0) &= \hat{u}^0, \\
\hat{u} &= 0 \quad \text{on } \partial\tilde{\Omega},
\end{aligned}$$

for $\tilde{\Omega}$ a bounded domain in \mathbb{R}^m with $m \geq 3$, and $\mathcal{L}(\hat{u})$ a second order elliptic operator with sufficiently smooth coefficients, provided that the space-time function F and \hat{u}^0 are sufficiently smooth also.

3.3. Lower bound

We multiply the equation of the transformed problem (3.1) with some function w in $H_0^1((0, T) \times \hat{\Omega})$ to be specified in the sequel and integrate in space-time so that

$$((\partial_t u, w))_{\hat{G}^n} = \int_{t^n}^{t^{n+1}} \mathcal{B}(\hat{u}, w) dt + ((\tilde{D}\hat{u}_y + \tilde{E}\hat{u}_\theta + \beta_1\hat{u}, w))_{\hat{G}^n} + ((\hat{f}, w))_{\hat{G}^n}.$$

Writing $\hat{u} = \hat{u} - u_h + u_h = \varepsilon + u_h$ we get

$$\begin{aligned} ((\partial_t \varepsilon, w))_{\hat{G}^n} &= \int_{t^n}^{t^{n+1}} \mathcal{B}(\varepsilon, w) dt - ((\tilde{D}\varepsilon_y + \tilde{E}\varepsilon_\theta + \beta_1\varepsilon, w))_{\hat{G}^n} \\ &= \sum_{K \in T_h^n} ((-\partial_t u_h + \hat{f} + Au_{hyy} + Bu_{hy\theta} + Cu_{h\theta\theta} + Du_{hy} + Eu_{h\theta} + \beta_1 u_h, w))_{I_n \times K} \\ &\quad - \int_{I_n} \sum_{K \in T_h^n} \sum_{\ell \in E_K^n} \int_{\ell} w [\mathcal{M} \nabla u_h \cdot \mathbf{n}]_{\ell} ds dt. \end{aligned} \quad (3.22)$$

Following the arguments of [8, 36, 43] and assuming that the meshes space-time meshes are regular with a regularity constant β_{reg} . Furthermore, we let

$$A_h := -\partial_t u_h + \hat{f} + Au_{hyy} + Bu_{hy\theta} + Cu_{h\theta\theta} + Du_{hy} + Eu_{h\theta} + \beta_1 u_h, \quad B_h := -[\mathcal{M} \nabla u_h \cdot \mathbf{n}]_{\ell}.$$

Then, there exists a constant c , depending on β_{reg} , and a function $w \in H_0^1((0, T) \times \hat{\Omega})$ such that for all $n = 0, \dots, N-1$ and for all $K \in T_h^n$ it holds that (cf. [36] pp. 225, 229),

$$w(\cdot, t^n) = 0, \quad (3.23)$$

$$|K| \int_{t^n}^{t^{n+1}} \|\Pi_K^n(A_h)\|_{L^2(K)}^2 dt = \int_{t^n}^{t^{n+1}} \int_K \Pi_K^n(A_h) w dx dt, \quad (3.24)$$

$$|\ell| \int_{t^n}^{t^{n+1}} \|\Pi_{\ell}^n(B_h)\|_{L^2(\ell)}^2 dt = \int_{t^n}^{t^{n+1}} \int_{\ell} \Pi_{\ell}^n(B_h) w ds dt \quad \forall \ell \in E_K^n, \quad (3.25)$$

$$\begin{aligned} |w|_{1, (I_n \times K)}^2 &\leq c \int_{t^n}^{t^{n+1}} \left\{ |K| \|\Pi_K^n(A_h)\|_{L^2(K)}^2 \right. \\ &\quad \left. + \sum_{\ell \in E_K^n} |\ell| \|\Pi_{\ell}^n(B_h)\|_{L^2(\ell)}^2 \right\} dt, \end{aligned} \quad (3.26)$$

where

$$\begin{aligned} \Pi_k^n(v) &:= \frac{1}{(t^{n+1} - t^n)|K|} \int_{t^n}^{t^{n+1}} \int_K v dx dt, \\ \Pi_{\ell}^n(v) &:= \frac{1}{(t^{n+1} - t^n)|\ell|} \int_{t^n}^{t^{n+1}} \int_{\ell} v ds dt. \end{aligned}$$

Here, $|w|_{1, (I_n \times K)}$ denotes the space-time seminorm. Let us now define

$$\eta_{lb}^n := \left[\int_{t^n}^{t^{n+1}} \left\{ |K| \|\Pi_K^n(A_h)\|_{L^2(K)}^2 + \sum_{\ell \in E_K^n} |\ell| \|\Pi_{\ell}^n(B_h)\|_{L^2(\ell)}^2 \right\} dt \right]^{1/2},$$

and note that from (3.26) we have that $|w|_{1,(I_n \times K)}^2 \leq c(\eta_{lb}^n)^2$. We then integrate by parts in (3.22) and use $w(\cdot, t^i) = 0$ to obtain

$$-((\varepsilon, w_t))_{\hat{G}^n} - \int_{t^n}^{t^{n+1}} \mathcal{B}(\varepsilon, w) dt - ((\tilde{D}\varepsilon_y + \tilde{E}\varepsilon_\theta + \beta_1\varepsilon, w))_{\hat{G}^n} = \sum_{K \in T_h^n} (\eta_{lb}^n)^2. \quad (3.27)$$

Using (3.26) we arrive at

$$\sum_{i=0}^{n-1} \sum_{K \in T_h^i} (\eta_{lb}^i)^2 \leq c \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \left\{ \|\varepsilon\|_K^2 + \|\nabla \varepsilon\|_K^2 \right\}. \quad (3.28)$$

These estimates readily provide lower bounds for the *a posteriori* estimate:

Theorem 3.7. *The error $\varepsilon = \hat{u} - u_h$ satisfies*

$$\frac{1}{c} \sum_{i=0}^{n-1} \sum_{K \in T_h^i} (\eta_{lb}^i)^2 \leq \sum_{i=0}^{n-1} \sum_{K \in T_h^i} \left\{ \|\varepsilon\|_K^2 + \|\nabla \varepsilon\|_K^2 \right\}, \quad (3.29)$$

where

$$(\eta_{lb}^i)^2 := \int_{t^i}^{t^{i+1}} \left\{ |K| \|\Pi_K^i(A_h)\|_{L^2(K)}^2 + \sum_{\ell \in E_K^i} |\ell| \|\Pi_\ell^i(B_h)\|_{L^2(\ell)}^2 \right\} dt, \quad (3.30)$$

and

$$\begin{aligned} A_h &:= -\partial_t u_h + \hat{f} + A u_{hy} + B u_{h\theta} + C u_{h\theta\theta} + D u_{hy} + E u_{h\theta} + \beta_1 u_h, \\ B_h &:= -[\mathcal{M} \nabla u_h \cdot \mathbf{n}]_\ell. \end{aligned} \quad (3.31)$$

We note that the lower bound, which is, of course, computable, is derived using projection to space-time constants. Furthermore, note that both Theorems 3.6 and 3.7 provide lower and upper *a posteriori* bounds, respectively, for the quantity

$$\sum_{i=0}^{n-1} \sum_{K \in T_h^i} \left\{ \|\varepsilon\|_K^2 + \|\nabla \varepsilon\|_K^2 \right\}.$$

4. NUMERICAL EXPERIMENTS

We applied our method for the problem (1.2) in dimensions $m = 2$ in space; recall that the space-time non-cylindrical domain S_T has the form

$$S_T = \{(x, t) : x \in \Omega(t), 0 < t < T\}.$$

We considered the case where

$$\Omega(t) := \{(z, \theta) \in \mathbb{R}^2 : z \in (0, s(t, \theta)), \theta \in (\theta_1, \theta_2)\},$$

for $t \in [0, T]$ and s a smooth positive function. We took

$$s(t, \theta) = 1 + 0.25 \sin(\pi\theta)t, \quad 0 \leq t \leq 1, \quad 0 < \theta < 1,$$

and defined

$$u(z, \theta, t) = e^{t/2} \frac{z\theta}{s^2(t, \theta)} \left(1 - \frac{z}{s(t, \theta)} \right)^3 \left(1 - \frac{\theta}{s(t, \theta)} \right), \quad (4.1)$$

TABLE 1. Estimated and exact errors for problem (1.2) with $m = 2$ and exact solution (4.1).

Grid size	h	k	\mathcal{E}	$\mathcal{E}_{\text{est}}/\mathcal{E}$
10×10	0.1	0.025	0.631744	3.26
20×20	0.05	0.025	0.327328	2.18
40×40	0.025	0.025	0.174575	1.61
80×80	0.0125	0.025	0.098069	1.94
10×10	0.1	0.0125	0.425477	2.45
20×20	0.05	0.0125	0.216215	1.88
40×40	0.025	0.0125	0.108654	1.46
80×80	0.0125	0.0125	0.054401	1.17

as the exact solution of (1.2).

We applied the proposed transformation to obtain an equivalent problem of the form (2.16). The new initial and boundary value problem, now posed on a cylindrical space-time domain, was approximated by the discontinuous Galerkin scheme (3.3). This scheme was implemented in a double precision Fortran/C++ code, where a uniform rectangular grid was used in space. To verify the theoretical result of Theorem 3.6 for the *a posteriori* upper bound, we computed the true error

$$\mathcal{E} = \sum_{i=0}^{N-1} \sum_{K \in \mathcal{T}_h^i} \|\nabla \varepsilon\|_K^2 + \sum_{i=0}^{N-1} \sum_{K \in \mathcal{T}_h^i} \|\varepsilon\|_K^2 + |\varepsilon^N|_{\Omega}^2 + \sum_{i=0}^{N-1} |u_h^{i+0} - u_h^i|_{\Omega}^2,$$

and the estimated error, \mathcal{E}_{est} , where we retained only the residual term and the 4th, 5th and 6th terms appearing in the right-hand side of the estimate of Theorem 3.6. We took $q = 3$ and used a range of space-time discretization parameters

$$h = \max_{0 \leq n \leq N} \mathbf{h}_n \quad k = \max_{0 \leq n \leq N} \mathbf{k}_n.$$

We performed two sets of experiments. In the first set we took $k = 0.025$ and varied the space discretisation parameter from $h = 0.1$ to $h = 0.0125$. In the second set of experiments we took $k = 0.0125$ and used the same values of the space discretisation parameter. We summarized our results in Table 1. The last column of Table 1 lists the *effectivity index*, computed as the ratio of the estimated error to the true error.

APPENDIX A.

This section presents some bounds concerning the local trace and local L^∞ Sobolev inequalities, used throughout our proofs. Proofs are furnished for bounded, multi-dimensional domains with the constants depending on the domains' diameters, so that the results can be applied locally on each element of our partition.

We begin by the trace inequality.

Theorem A.1. *Let \mathcal{A} be a bounded domain in \mathbb{R}^n with Lipschitz boundary. There exists a positive constant c , independent of \mathcal{A} , such that*

$$\|u\|_{L^2(\partial\mathcal{A})}^2 \leq c\|u\|_{L^2(\mathcal{A})}\|u\|_{H^1(\mathcal{A})} + c(\text{diam}\mathcal{A})^{-1}\|u\|_{L^2(\mathcal{A})}^2, \quad \text{for any } u \in H^1(\mathcal{A}). \quad (\text{A.1})$$

Proof. We use the general trace inequality (cf. [15])

$$\|v\|_{L^p(\partial\hat{\mathcal{A}})} \leq c\|v\|_{L^p(\hat{\mathcal{A}})}^{1-1/p}\|v\|_{W_p^1(\hat{\mathcal{A}})}^{1/p}$$

with $p = 2$ and

$$\hat{\mathcal{A}} := \left\{ y \in \mathbb{R}^n : \exists x \in \mathcal{A} : y = \frac{x}{\text{diam}\mathcal{A}} \right\},$$

so that

$$\|v\|_{L^2(\partial\hat{\mathcal{A}})}^2 \leq c\|v\|_{L^2(\hat{\mathcal{A}})}\|v\|_{H^1(\hat{\mathcal{A}})},$$

where c is independent of $\hat{\mathcal{A}}$, since the measure of $\hat{\mathcal{A}}$ is of order $\mathcal{O}(1)$. In addition, using the change of variables $y = \frac{x}{\text{diam}\mathcal{A}}$, $x \in \mathcal{A}$ and $\hat{u}(y) = u(x)$, we have, for any integer k , that

$$|\hat{u}|_{H^k(\hat{\mathcal{A}})} = (\text{diam}\mathcal{A})^{k-\frac{n}{2}}|u|_{H^k(\mathcal{A})}. \quad (\text{A.2})$$

Observing now the scaling for the surface area S_{n-1} of a ball in \mathbb{R}^n of radius ρ , that is

$$\int_{S_{n-1}} 1 \, ds = nc_n \rho^{n-1},$$

we obtain, using (A.2),

$$\begin{aligned} \int_{\partial\mathcal{A}} |u|^2 \, ds &\leq c(\text{diam}\mathcal{A})^{n-1} \int_{\partial\hat{\mathcal{A}}} |\hat{u}|^2 \, ds \leq c\|\hat{u}\|_{L^2(\hat{\mathcal{A}})}\|\hat{u}\|_{H^1(\hat{\mathcal{A}})}(\text{diam}\mathcal{A})^{n-1} \\ &\leq c(\text{diam}\mathcal{A})^{0-\frac{n}{2}}\|u\|_{L^2(\mathcal{A})}\left[(\text{diam}\mathcal{A})^{1-\frac{n}{2}}|u|_{H^1(\mathcal{A})} + (\text{diam}\mathcal{A})^{0-\frac{n}{2}}|u|_{L^2(\mathcal{A})}\right](\text{diam}\mathcal{A})^{n-1} \\ &\leq c\|u\|_{L^2(\mathcal{A})}\|u\|_{H^1(\mathcal{A})} + c(\text{diam}\mathcal{A})^{-1}\|u\|_{L^2(\mathcal{A})}^2. \end{aligned}$$

for some constant c , independent of \mathcal{A} . \square

The next result presents the L^∞ Sobolev inequality for bounded domains of \mathbb{R}^n , for $n = 1, 2$, with constants depending on the diameter of the domain. The estimate for $n = 2$ is not used in the proofs of this paper, but we include it for the sake of completeness.

Theorem A.2. *Let \mathcal{A} be a bounded domain in \mathbb{R}^n , $n = 1, 2$, with Lipschitz boundary, if $n = 2$. Then, there exists a positive constant c , independent of \mathcal{A} , such that*

$$\|u\|_{L^\infty(\mathcal{A})}^2 \leq c(\text{diam}\mathcal{A})^{-1}\|u\|_{L^2(\mathcal{A})}^2 + c(\text{diam}\mathcal{A})|u|_{H^1(\mathcal{A})}^2, \quad \text{if } n = 1, \quad (\text{A.3})$$

and

$$\|u\|_{L^\infty(\mathcal{A})}^2 \leq c(\text{diam}\mathcal{A})^{-2}\|u\|_{L^2(\mathcal{A})}^2 + c|u|_{H^1(\mathcal{A})}^2 + c(\text{diam}\mathcal{A})^2|u|_{H^2(\mathcal{A})}^2, \quad \text{if } n = 2. \quad (\text{A.4})$$

Proof. Using the notation of the previous theorem, we note first that

$$\|u\|_{L^\infty(\mathcal{A})} = \|\hat{u}\|_{L^\infty(\hat{\mathcal{A}})}. \quad (\text{A.5})$$

If $n = 1$, the Sobolev inequality and (A.2) give [15]

$$\begin{aligned} \|\hat{u}\|_{L^\infty(\hat{\mathcal{A}})}^2 &\leq c\|\hat{u}\|_{H^1(\hat{\mathcal{A}})}^2 = c\|\hat{u}\|_{L^2(\hat{\mathcal{A}})}^2 + c|\hat{u}|_{H^1(\hat{\mathcal{A}})}^2 \\ &= c(\text{diam}\mathcal{A})^{2(0-1/2)}\|u\|_{L^2(\mathcal{A})}^2 + c(\text{diam}\mathcal{A})^{2(1-1/2)}|u|_{H^1(\mathcal{A})}^2, \end{aligned}$$

for some positive constant c independent from \mathcal{A} . Hence, by (A.5) and the above relation, the inequality (A.3) follows. Let us now consider the case $n = 2$. The Sobolev inequality combined with (A.2) give

$$\begin{aligned} \|\hat{u}\|_{L^\infty(\hat{\mathcal{A}})}^2 &\leq c\|\hat{u}\|_{H^2(\hat{\mathcal{A}})}^2 = c\|\hat{u}\|_{L^2(\hat{\mathcal{A}})}^2 + c|\hat{u}|_{H^2(\hat{\mathcal{A}})}^2 \\ &= c(\text{diam}\mathcal{A})^{2(0-2/2)}\|u\|_{L^2(\mathcal{A})}^2 + c(\text{diam}\mathcal{A})^{2(1-2/2)}|u|_{H^1(\mathcal{A})}^2 + c(\text{diam}\mathcal{A})^{2(2-2/2)}|u|_{H^2(\mathcal{A})}^2, \end{aligned}$$

so (A.4) follows easily by using the above relation and (A.5). \square

Acknowledgements. The authors would like to thank Prof. Ch. Makridakis for stimulating discussions, and the two anonymous referees for their valuable comments.

REFERENCES

- [1] G. Akrivis, V.A. Dougalis and G.E. Zouraris, Finite difference schemes for the “Parabolic” equation in a variable depth environment with a rigid bottom boundary condition. *SIAM J. Numer. Anal.* **39** (2001) 539–565.
- [2] G. Akrivis and C. Makridakis, Galerkin time-stepping methods for nonlinear parabolic equations. *ESAIM: M2AN* **38** (2004) 261–289.
- [3] R. Alexander, Diagonally Implicit Runge-Kutta methods for stiff O.D.E.’s. *SINUM* **14** (1977) 1006–1021.
- [4] D.C. Antonopoulou, *Theory and numerical analysis of parabolic approximations*. Ph.D. thesis, Department of Mathematics, University of Athens, Greece (2006).
- [5] D.C. Antonopoulou, V.A. Dougalis and G.E. Zouraris, Galerkin methods for Parabolic and Schrödinger Equations with dynamical boundary conditions and applications to underwater acoustics. *SIAM J. Numer. Anal.* **47** (2009) 2752–2781.
- [6] D.C. Antonopoulou, G.D. Karali, M. Plexousakis and G.E. Zouraris, Crank–Nicolson finite element discretizations for a 2D linear Schrödinger-type equation posed in a noncylindrical domain. *Math. Comput.* **84** (2015) 1571–1598.
- [7] D.C. Antonopoulou and M. Plexousakis, Discontinuous Galerkin methods for the linear Schrödinger equation in non-cylindrical domains. *Numer. Math.* **115** (2010) 585–608.
- [8] I. Babuska, R. Duran and R. Rodriguez, Analysis of the efficiency of an a posteriori error estimator for linear triangular finite elements. *SIAM J. Numer. Anal.* **29** (1992) 947–964.
- [9] M.J. Baines, Moving Finite Elements, Monographs on Numerical Analysis. The Clarendon Press, Oxford University Press, Oxford Science Publications, New York, NY (1994).
- [10] J. Baranger and H. El-Amri, Estimateurs a posteriori d’ erreur pour le calcul adaptif d’écoulements quasi-newtoniens. *RAIRO Model. Math. Anal. Numer.* **25** (1991) 31–48.
- [11] C. Bernardi, Optimal finite element interpolation on curved domains. *SIAM J. Numer. Anal.* **26** (1989) 1212–1240.
- [12] K.S. Bey, A. Patra and J.T. Oden, hp-version discontinuous Galerkin methods for hyperbolic conservation laws: a parallel strategy. *Internat. J. Numer. Methods Eng.* **38** (1995) 3889–3908.
- [13] K.S. Bey, A. Patra and J.T. Oden, hp-version discontinuous Galerkin methods for hyperbolic conservation laws. *Comput. Methods Appl. Mech. Eng.* **133** (1996) 259–286.
- [14] K.S. Bey, A. Patra and J.T. Oden, A parallel hp-adaptive discontinuous Galerkin method for hyperbolic conservation laws. *Appl. Numer. Math.* **20** (1996) 321–386.
- [15] S.C. Brenner and L. Ridgway Scott, The Mathematical Theory of Finite Element Methods. Springer-Verlag, New York, NY (1991).
- [16] P. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **9** (1975) 77–84.
- [17] B. Cockburn and C.W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.* **35** (1998) 2440–2463.
- [18] M. Crouzeix, Sur l’approximation des équations différentielles opérationnelles linéaires par des méthodes de Runge Kutta, Thèse, University Paris VI, Paris (1975).
- [19] L.C. Evans, Partial Differential Equations. American Mathematical Society, Providence, RI (1998).
- [20] G. Fichera, On a unified theory of boundary value problems for elliptic – parabolic equations of second order. In: Boundary Problems in Differential Equations. University of Wisconsin Press, Madison, WI (1960).
- [21] A. Hiltebrand and S. Mishra, Entropy stable shock capturing space-time discontinuous Galerkin schemes for systems of Conservation Laws. Research Report No. 2012-07, ETHZ (2012).
- [22] A. Hiltebrand, S. Mishra and C. Parés, Entropy-stable space-time DG schemes for non-conservative hyperbolic systems. *ESAIM: M2AN* **52** (2018) 995–1022.
- [23] P. Jamet, Numerical methods and existence theorems for parabolic differential equations whose coefficients are singular on the boundary. *Math. Comput.* **22** (1968) 721–743.
- [24] P. Jamet, Estimation d’erreur pour des éléments finis droits presque dégénérés. *Rev. Francaise Automat. Informat. Recherche Opérationnelle Sér. Rouge Anal. Numér.* **10** (1976) 43–61.
- [25] P. Jamet, Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM J. Numer. Anal.* **15** (1978) 913–928.
- [26] O. Karakashian and C. Makridakis, A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous galerkin method. *Math. Comput.* **67** (1998) 479–499.
- [27] O.A. Karakashian and F. Pascal, A posteriori error estimates for a Discontinuous Galerkin approximation of second-order Elliptic problems. *SIAM J. Numer. Anal.* **41** (2003) 2374–2399.
- [28] C. Klaij, J.J.W.V. der Vegt and H.V. der Ven, Space-time discontinuous Galerkin method for the compressible Navier–Stokes equations. *J. Comput. Phys.* **217** (2006) 589–611.
- [29] P. Lesaint and P.A. Raviart, On a finite element method for solving the neutron transport equation, edited by C. deBoor. In: Mathematical Aspects of Finite Elements in Partial Differential Equations. Academic Press, New York, NY (1974).
- [30] J.L. Lions, Sur les problèmes mixtes pour certains systèmes paraboliques dans des ouverts non cylindriques. *Ann. Inst. Fourier* **7** (1957) 143–182.
- [31] C. Makridakis and R.H. Nochetto, A posteriori error analysis for higher order dissipative methods for evolution problems. *Numer. Math.* **104** (2006) 489–514.
- [32] G. May and M. Zakerzadeh, On the convergence of space-time discontinuous galerkin schemes for scalar conservation laws. *SIAM J. Numer. Anal.* **54** (2016) 2452–2465.

- [33] A.L. Mignot, Methodes d'approximation des solutions de certains problemes aux limites lineaires. *Rend. Sem. Mat. Univ. Padova, XL* **40** (1968) 1–38.
- [34] K. Miller, Moving finite elements, II. *SINUM* **18** (1981) 1033–1057.
- [35] L. Pareschi and G. Russo, Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *J. Sci. Comput.* **25** (2005) 129–155.
- [36] M. Picasso, Adaptive finite elements for a linear parabolic problem. *Comput. Meth. Appl. Mech. Eng.* **167** (1998) 223–237.
- [37] W.H. Reed and T.R. Hill, Triangular mesh methods for the neutron transport equation. Tech. Report LA-UR-73-479. Los Alamos Scientific Laboratory (1973).
- [38] D. Schötzau and T.P. Wihler, A posteriori error estimation for hp-version time-stepping methods for parabolic partial differential equations. *Numer. Math.* **115** (2010) 475–509.
- [39] M. Tavelli and M. Dumbser, A staggered space-time discontinuous Galerkin method for the three-dimensional incompressible Navier–Stokes equations on unstructured tetrahedral meshes. *J. Comput. Phys.* **319** (2016) 294–323.
- [40] M. Tavelli and M. Dumbser, Arbitrary high order accurate space-time discontinuous Galerkin finite element schemes on staggered unstructured meshes for linear elasticity. *J. Comput. Phys.* **366** (2018) 386–414.
- [41] J.J.W. van der Vegt and H. van der Ven, Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flows I. general formulation. *J. Comp. Phys.* **182** (2002) 546–585.
- [42] H. van der Ven and J.J.W. van der Vegt, Space-time discontinuous Galerkin finite element method with dynamic grid motion for inviscid compressible flows II. efficient flux quadrature. *Comput. Methods Appl. Mech. Eng.* **191** (2002) 4747–4780.
- [43] R. Verfürth, error estimators for the Stokes equations. *Numer. Math.* **55** (1989) 309–325.